DOCUMENT RESUME

ED 036 899 EA 002 791

AUTHOR	Crockenberg, Vincent, Ed.
TITLE	American Educational Research Association Paper Abstracts, 1969 Annual Meeting.
INSTITUTION	American Educational Research Association, Washington, D.C.
PUB DATE	Feb 69
NOTE	279p.; Abstracts of papers presented at the Paper Sessions of the Annual Meeting of the AERA (Los Angeles, Calif., Feb. 5-8, 1969)
AVAILABLE FROM	American Educational Research Association, 1126 Sixteenth St., N.W., Washington, D.C. 20036 ($2.00)
EDRS PRICE	EDRS Price MF-$1.25 HC Not Available from EDRS.
DESCRIPTORS	*Abstracts, Academic Achievement, Cognitive Development, *Conference Reports, Curriculum Development, *Educational Research, Evaluation Techniques, Instructional Materials, Learning Processes, Measurement Techniques, *National Organizations, Problem Solving, *Research Methodology, Research Tools, Statistical Analysis, Teacher Behavior, Teacher Characteristics, Testing

ABSTRACT

This document contains abstracts of papers presented during the 93 sessions of the 1969 annual meeting of the American Educational Research Association. The papers use a wide variety of research approaches and techniques to investigate such topics as teacher characteristics and behavior, the learning process in various settings, curriculum development, student achievement, testing methods, cognitive development, instructional technology, individual differences, problem solving, and political and administrative influences on educational outcomes. A list of program participants is included. Related documents are EA 002 792, EA 002 793, and EA 002 794. (JH)

N-X

ED036899

American Educational Research Association

PAPER ABSTRACTS

Editor

Vincent Crockenberg

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY
Michael McCormick
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

Paper Sessions of the

1969 ANNUAL MEETING

Los Angeles
Biltmore and Statler Hilton Hotels
February 5-8

American Educational Research Association
1126 Sixteenth Street, N.W.
Washington, D.C. 20036

EA 002 791

Copyright © 1969
by
American Educational
Research Association
1126 Sixteenth Street, N.W.
Washington, D.C. 20036

$2.00

All rights reserved. No part of this book may be reproduced in any form, by mimeograph or any other means, without permission in writing from the American Educational Research Association.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

A.E.R.A. Division Program Committees
1968-1969

The following A.E.R.A. members served as evaluation teams for the contributed papers abstracted in this volume:

A.E.R.A. PROGRAM CHAIRMAN—H. Thomas James
CO-CHAIRMAN—Arthur P. Coladarci

Division

A—James A. Kelly, Chairman
  Frank Lutz
  William Monohan
  Jay D. Scribner
B—Kenneth J. Rehage, Chairman
  William M. Alexander
  George A. Beauchamp
  Arno A. Bellack
  Elliot W. Eisner
C—Joe L. Byers, Chairman
  Laurence D. Brown
  Robert C. Craig
  Robert E. Davidson
  Gary Davis
  Frank Farley
  Jim Gaite
  Charles Gaverick
  Robert Grinder

Division

  Howard Kight
  William W. Lynch
  Joan Prentice
  Thomas J. Shuell
  Lee S. Shulman
  Richard L. Turner
  Joanna Williams
  Joseph Wolff
D—Joan Bollenbacher, Chairman
  John Milholland
  William Turnbull
E—Gary R. Walz, Chairman
  John O. Crites
  Richard T. Knowles
  E. Paul Torrance
F—Charles Burgess, Chairman
  Merle L. Borrowman
  David Tyack

TABLE OF CONTENTS

Page
AERA DIVISION OF PROGRAM COMMITTEES iii

PAPER ABSTRACTS

SESSIONS

1.5 Psycholinguistics No. 1 1
1.9 Acquisition of Reading Skills No. 1 3
1.10 Interactions of Individual Difference and Task Variables 6
1.16 Contributed Papers 9
2.2 Problem Solving No. 1 12
2.3 Teacher Education No. 1 15
2.4 Development of Measurement Instruments 17
3.6 Computer Assisted Instruction No. 1 20
3.7 Cognitive Development 23
3.8 Problem Solving No. 2 25
3.13 Factor Analysis—A Technique and a Tool 28
3.14 Institutional Founders 30
4.3 Computer Assisted Instruction No. 2 31
4.5 Individual Differences No. 1 34
4.6 Empirical Research on Objective and Subjective Testing 37
5.5 Individual Differences No. 2 39
5.7 Learning from Discourse: Mathemagenics No. 1 42
5.10 Empirical Test Research 44
5.11 Counselor Characteristics and Counseling Innovations 47
6.3 Verbal Learning No. 1 49
6.4 Preschool Training Programs No. 1 52
6.7 Transformation of the School Revisited 54
7.1 Models for Curriculum Development 57
7.2 Conflict and Innovation 61
7.3 Political Socialization and Other Approaches to the Study of School Outcomes 64
7.4 Instructional Materials 66
7.5 Programmed Instruction 69
7.6 Psycholinguistics No. 2 72
7.7 Teacher Behavior 75
7.9 Studies of Achievement and Intelligence 78
7.12 Research on Testing 81
8.1 Studies of Organizational Climate 84
8.7 Acquisition of Reading Skills No. 2 88
8.8 Contributed Papers No. 1 91
8.9 Verbal Learning No. 2 101
8.10 Concept Learning No. 1 104
8.15 Research on Teaching 107
9.1 Preschool Training Programs No. 2 109
9.2 Micro-Teaching 112
9.3 Creativity No. 1 115
9.4 Approaches to Research Design 117
10.2 Quantitative Techniques for Educational Planning and Evaluation 120
10.3 Studies for Pre-School Programs 122
10.7 Personnel Administration 125
10.8 Teacher Characteristics No. 1 127
10.13 Statistical Methods 129
11.4 Social Studies Instruction 132

CONTENTS

Page
11.5 Verbal Learning No. 3 135
11.6 Administration and Research in Higher Education 137
11.7 Mathematics Instruction No. 1 140
11.8 Statistical Methods: Analysis of Variance and Covariance 142
12.4 Interaction Analysis No. 1 145
12.5 Acquisition of Reading Skills No. 3 147
12.7 Feedback and Reinforcement No. 1 150
12.11 The Impact of Personal and Environment Factors on School Related Performance 152
12.12 Experimental Design and Analysis 155
12.13 Research on Student Personnel Services and Students 159
13.5 Cognitive Development: Multiple Classification 168
13.6 Individual Differences No. 3 170
14.2 Use of Objectives in Curriculum Development 174
14.5 Cognitive Development: Piagetian Conservation No. 1 176
14.6 The Local Socio-Economic Context of Educational Administration 178
14.7 Contributed Papers 182
14.8 Feedback and Reinforcement No. 2 187
14.9 Mathematics Instruction No. 2 189
14.14 Measurement of Young Children 191
15.2 Social Class, Race and Achievement 193
15.3 Concept Learning No. 2 196
15.5 Learning from Discourse: Proaction and Retroaction Effects 199
16.3 Achievement in Primary Grades 201
16.5 Contributed Papers No. 2 203
16.6 Teacher Education No. 2 206
16.7 Teacher Characteristics No. 2 208
16.10 School-Environment Influence of Child Behavior 211
16.11 Statistical Method and Theory 214
17.1 Evaluation in Curriculum Development 220
17.2 Motivation 223
17.3 Politics of Education at the State and National Level 226
17.4 Contributed Papers 229
17.5 Social Class and Language 232
17.8 Research on Teacher Behavior 234
17.9 Creativity No. 2 237
18.5 Interaction Analysis No. 2 240
18.6 Special Education 243
18.7 Meaningful Verbal Learning 245
18.10 Research on Tests 247
18.11 Studies Relating to Verbal Performance 250
19.1 Cognitive Development: Piagetian Conservation No. 2 253
19.2 Bureaucracy and Innovation 255
19.3 Concept Learning No. 3 258
19.4 Learning from Discourse: Mathemagenics No. 2 261
19.6 Multivariate Analysis 263
19.8 Application of Student Response Measures 266
PROGRAM PARTICIPANTS 269
OFFICERS OF AERA ....Back Cover

PSYCHOLINGUISTICS NO. 1

Joan Prentice, Indiana University Chairman

A STUDY COMPARING GLOBAL QUALITY WITH SYNTACTIC MATURITY IN THE WRITTEN COMPOSITION OF SECOND AND THIRD GRADE STUDENTS, *Edieann Biesbrock*, University of Minnesota; *L. Ramon Veal*, University of Georgia

Sixty papers were selected from compositions produced by approximately 1,000 second and third graders. The papers represented a stratified, *one*-through-*seven* distribution of levels of quality on a global essay scale developed at the University of Georgia. Each essay had been rated by four raters with perfect agreement. The papers were divided into T-units, i.e., one main clause and its modifiers, and the following counts were made: the number of garbles per paper, the number of T-units per paper, the number of subordinate clauses per paper, the number of all (total) clauses per paper, the ratio of clauses per T-unit, and mean T-unit length.

Observed correlations between global essay ratings and the number of T-units (.75) and the ratio of clauses per T-unit (.62) were moderately high. However, correlations between global essay ratings and the number of subordinate clauses (.71) and the number of all clauses (.80) were comparable. Among the seven levels of rated quality for this sample of children, these two measures were as reliable as T-unit counts. The correlation between global essay ratings and mean T-unit length was lower (.48). Although Hunt (1965) used mean T-unit length to distinguish between fourth, eighth, and twelfth graders, it was not, in this study, a reliable method of distinguishing between levels of quality within a specific age group. Garbles also proved to be no indication of quality.

All measures, other than the number of garbles and mean T-unit length, related directly to the global rating of quality. In summary, the direction of the differences between the *one*-through-*seven* levels on the global essay scale suggests that possibly both fluency and syntactic complexity are related to quality in composition at this early elementary level.

CHILDREN'S COMPREHENSION OF BETWEEN AND WITHIN SENTENCE SYNTACTIC STRUCTURES, *John R. Bormuth*, University of Chicago; *John Manning, Julian Carr, David Pearson*, University of Minnesota

These experiments were (a) to determine if selected linguistic structures differed in comprehension difficulty, (b) to determine if the question types used differed in difficulty, and (c) to gain a preliminary estimate of children's ability to comprehend common linguistic constructions.

At the intrasentence syntax level, 24 structures were studied, where structures were defined by the transformations performed on simple matrix sentences. At the anaphoric and intersentence syntax levels 14 and 16, respectively, structures were studied.

Two instances, a sentence or sentence pair, were constructed to represent each structure and each was embedded in a different four sentence paragraph of connected prose. Vocabulary was selected from the *Dale List of 3000 Easy Words* and no sentence contained more than one transformation. A rote, transform, semantic substitute, and compound WH question deleting the modifier of the construction tested was derived for each structure and used to make four test forms for each paragraph, the paragraph with one question below it being printed on one sheet of paper.

Fourth grade students were given one question form over each of the paragraphs with question forms randomly assigned to subjects. Sixty subjects answered each question under untimed conditions. Responses were scored using definitions of item responses which yielded very high interscores agreement. The proportion answering each question correctly was analyzed in a structure-by-question type design. Both main effects were significant for all three levels of structures indicating that instruction in the comprehension of these structures and questions should be sequenced by difficulty. The grand means were .58, .77, and .75 over the intersentence, anaphora, and intrasentence structures, respectively. This was viewed as poor performance in view of the extremely simple language tasks presented, suggesting that children's reading comprehension instruction is inadequate to permit them to approach mastery on verbally presented instruction.

COMPETENCE VS. PERFORMANCE IN YOUNG CHILDREN'S USE OF COMPLEX LINGUISTIC STRUCTURES, *Juanita Bryson* and *Carolyn Stern*, University of California, Los Angeles

Comparatives are a feature of elaborated language in which disadvantaged children, especially those from Mexican-American homes, are presumed to be deficient. The major hypothesis of the present study was that a relatively simple program could teach children from such language--different backgrounds to produce comparatives. A secondary hypothesis compared the effectiveness of two instructional procedures: (1) allowing the children to induce the rule for forming comparatives, and (2) presenting the simple verbal rule: add -er.

Twelve children, 50 to 60 months old, were shown simple black-and-white line drawings of paired pictures, each pair visually representing a comparison (e.g. big-bigger; fat-fatter) accompanied by appropriate verbal commentary on magnetic tape. A pretest, three 15-minute lessons, a posttest, and a retention test five months later were given to the same group of children.

On all tests, 2 types of responses were required: selection and verbal production. The selection task (on which most children scored close to ceiling) was used to determine whether the children understood the concept of a comparative relationship even if they did not have the appropriate word. The second task was more demanding, requiring that the child actually produce the comparative form. While the difference between the mean pre- and posttest scores (1.5 and 21.8, respectively) on this section was statistically significant (.01), no difference attributable to instructional treatment was found. The retention mean (20.0) was not measurably lower than the posttest, indicating an unexpectedly stable change in verbal behavior. That this was not due to maturation alone was inferred, since untrained children, from the same population, given the retention test performed at the same level (1.5) as had their peers on the pretest.

Because of the ease with which children produced and retained the ability to form comparatives, it might be argued that this study supports the "competence" hypothesis held by linguists. However, it seems clear that "competence" does not become "performance" without an instructional program.

MASS AND COUNT NOUN RESPONSES OF YOUNG CHILDREN, *Evelyn Hatch*, Southwest Regional Laboratory for Educational Research and Development

Forty-one kindergarteners and 23 pre-second graders were given a two-part task to determine their ability to use the syntactic cues for mass and count nouns.

Each task used ten count nouns and ten mass nouns. In Task 1, a participation story, the child responded to a "How much + mass noun?" or to a "How many + count noun?" question while looking at a pictured item. Task 2 required him to look at a pictured item and to ask "How much + noun?" or "How many + noun?"

Responses to Task 1 were scored for accuracy according to determiner agreement; those to Task 2, according to much/many agreement with the noun and number agreement.

Results suggest that children entering a reading program have not mastered the distinction between mass and count nouns. Their ability to use agreement cues of syntax appropriately has improved by age 7, although their handling of syntactic cues for mass nouns continues to be a problem even at this age.

Session 1.9 **Division C**

ACQUISITION OF READING SKILLS NO. 1

Wayne R. Otto, University of Wisconsin

SIMULTANEOUS AND SUCCESSIVE DISCRIMINATION LEARNING AS A FUNCTION OF STIMULUS SIMILARITY, *Margaret D. Ackerman* and *Joanna P. Williams*, University of Pennsylvania

This experiment was designed to investigate the relative difficulty of simultaneous and successive discrimination of letters under varying degrees of stimulus similarity. Lipsitt (1961) did a series of discrimination experiments, using fourth graders as subjects and colors a stimuli, which yielded contradictory results. Presumably, his subjects had distinct labels available for each of the color stimuli, and this may have attenuated the effects of stimulus similarity. For this reason, first graders were run in the present study, and the stimuli were letters which the Ss could not consistently label.

In the first experiment, the similar stimuli were *b* and *d*, and the dissimilar stimuli were *b* and *s*. In the simultaneous discrimination problem, S was consistently reinforced for responding to one of the stimuli regardless of its spatial position. In the successive discrimination problem, only one stimulus was presented on each trial and the S learned to make a differential positional response to each stimulus. Experiment 2 included colors (red and blue) as an additional class of stimuli.

The results show that neither the method of stimulus presentation (simultaneous or successive) nor the similarity of the stimuli, was a significant factor in either experiment. There is some indication that with similar letter stimuli, the successive problem was less difficult than the simultaneous; but with dissimilar letter stimuli, the successive problem was more difficult. The results are discussed in terms of discrimination theory and in terms of implications for educational practice.

AN EXPERIMENTAL STUDY OF VISUAL PERCEPTUAL TRAINING AND READINESS SCORES WITH CERTAIN FIRST-GRADE CHILDREN, *James DeWitt Cowles*, College of William and Mary

The purpose of the study was to investigate the efficacy of specific visual perceptual training on readiness scores. The null hypothesis tested was that there were no significant differences among the three groups.

An Experimental and two Control Groups were formed: the Experimental Group received visual perceptual training drawn from the *Beginning Pictures and Patterns* portion of *The Development Program in Visual Perception* by Marianne Frostig; the Instructional Control Group received listening activities; and the Control Group received no specific treatment with respect to gains in readiness as measured by the *Metropolitan Readiness Tests,* Forms A and B.

The total first-grade population (N=257) in two predominantly black elementary schools in Tuscaloosa, Alabama were used for selection purposes. The *Metropolitan Readiness Tests, Form B* was administered to the total population. The sample, selected on the basis of pre-test scores, consisted of 81 pupils chosen by stratified randomization.

The subjects in the Experimental Group (N=27) and the Instructional Control Group (N=27) were removed from their regular classrooms for experimental treatment which was administered three times weekly for a period of nine weeks. Twenty-seven lessons of 30 minutes duration were taught. The subjects were post-tested with the *Metropolitan Readiness Tests, Form A.*

Differences between pre- and post-test measures, as well as gains from pre- to post-test, were analyzed by means of the analysis of variance (Type I Design) for the three groups.

All obtained values of *F* resulting from the Type I analysis were statistically significant beyond the .05 level of significance. The one way analysis of variance technique was applied to the post-test scores taken two at a time. The values of *F* obtained between the Experimental Group and the Instructional Control Group and between the Experimental Group and the Control Group were statistically significant beyond the .001 level of significance. The value of *F* obtained between the Instructional Control Group and the Control Group was not statistically significant at the .05 level of significance.

It was concluded that the significant differences obtained between the Experimental Group and the Control Group, indicated the efficacy of visual perceptual instruction drawn from *The Developmental Program in Visual Perception* on readiness, as measured by the *Metropolitan Readiness Tests.* That no significant differences were found between the Instructional Control Group and the Control Group indicated that the effect of visual perceptual treatment, rather than interaction with the investigator, was operative.

The findings strongly suggest that *The Developmental Program in Visual Perception* can be of merit and is worthy of further empirical research.

ERRORLESS DISCRIMINATION LEARNING OF SIMILAR LETTERS OF THE ALPHABET, *R. J. Karraker*, University of Missouri, Kansas City, *Larry A. Doke*, Kansas University.

The "conditioning and extinction" theory of discrimination learning advanced by Hull and Spence has recently been challenged in a series of articles by Terrace. Terrace has demonstrated that discrimination learning can occur without errors. The discrimination was more accurate, less "emotional," and less susceptible to the influence of drugs.

Terrace accomplished this by fading the brightness, duration of presentation, and wave length of the stimuli. Goldiamond and Moore demonstrated the phenomena in a matching to sample task using young children, but the tasks were not directly similar to the types of discriminations children learn in school settings.

The lower case letters 'b' and 'd' were taught to 64 kindergarten children. Ss were seated in an enclosed chamber and presentation of stimuli was accomplished by rear-projecting slides which were articulated to a tape recorder for auditory stimuli. E provided knowledge of results through an intercom.

Consisting of three sessions of approximately 80 presentations each, the experiment had two variables with two values for each variable. The first variable was type of acquisition stimuli, with two values—Progressive and Constant. Progressive was defined as fading size, color, pictorial prompts, and duration of presentation. Constant was defined as no fading.

The second variable was time of introduction of the second component of the discrimination, with two values—Early and Late. Thus, there were four groups of 16 Ss each: Early-Progressive, Late-Progressive, Early-Constant, and Late-Constant.

A treatment (E-P) by levels (E-L) analysis of covariance with mental age as the covariate revealed a significant effect for levels ($p=.01$) but not for treatments. Data are also presented on learning acquisition as a function of the four treatments.

RECOGNITION OF FLASHED WORDS BY CHILDREN, *S. Jay Samuels*, University of Minnesota

The purpose of this study was to find what effect word associations have on speed of word recognition. An identical study using adults disclosed that when associated adjective-noun word pairs were presented together, recognition speed was significantly faster to the noun than when a noun was presented by itself. Conversely, when adjective-noun word pairs were presented which were not associates of each other, recognition speed was significantly slower in comparison to the noun by itself.

A 5 x 5 repeated-measure Latin square design was used with 20 fourth-graders. First, the subject was familiarized with the words which would be shown in the tachistoscope. This was done by having him read single words (nouns) or associated adjective-noun word-pairs aloud from cards. Then the same single words and word-pairs were tachistoscopically presented, either in the same context or in a different context. Recognition speed was timed for the noun only.

The results indicated that word recognition strategies of children are different from those of adults. For example, with children speed of recognition was as fast when a noun was presented by itself as when it was preceded by its associate. Another difference was found when associated word-pairs were flashed which had not been presented during familiarization. With adults there was no difference in recognition speed between familiarized and non-familiarized word-associates, but there was with children. For adults as well as for children, recognition speed was severely retarded when non-associated word-pairs were flashed. These results along with others will be discussed in terms of strategies children use in word recognition.

Session 1.10 **Division C**

INTERACTIONS OF INDIVIDUAL DIFFERENCE AND TASK VARIABLES

Robert D. Tarte, University of Michigan, Chairman

COGNITIVE-DOMAIN ACHIEVEMENT IN SCIENCE, A FUNCTION OF METHOD OF PRESENTATION, INTELLECTUAL CAPACITY, AND PERSONALITY ADJUSTMENT, *Willim S. Donaldson,* Department of Public Instruction, State of Pennsylvania

Empirical investigations of the learning process have shown that achievement potential is better utilized when instruction is individualized. The purpose of this study was to examine several relationships between method of instruction and selected student characteristics. It was hypothesized that achievement could be increased when unique student-types were compatibly matched with alternate instructional presentations. Definition of student-type was based on level of intellectual capacity and level of personality adjustment.

Fourth-grade pupils were provided linearly-structured programmed texts for elementary-level science. Included were experiments to be performed using simple apparatus. All students used identical texts but the experiments were completed via three alternate methods. Each child was provided a text into which one of the three learning presentations had been structured.

Level of intellectual capacity and level of personality adjustment were highly correlated with achievement. Within levels of intellectual capacity the highest level of personality adjustment had the highest achievement, the lowest level of personality adjustment had the lowest achievement. The reverse was also true: within levels of personality adjustment the highest intellectual capacity exhibited the greatest acquisition of cognitive-type learning. Also, within categories such as high-level intellectual capacity and low-level personality adjustment, it was found that additional partitioning by instructional method produced an F-ratio significant at the 0.005 level.

Specific student-types learned more when the presentation was of a specific design. Further, it was demonstrated that instructions could be individualized so as to provide alternate routes for learners, the implementation of which resulted in increased achievement.

SOME SITUATIONAL AND EXTRINSICLY BASED DIMENSIONS OF THE CONSTRUCT "MOTIVATIONAL DISTURBANCE" AND ITS EFFECT ON STUDENT DECISION-MAKING IN AN INSTRUCTIONAL CONTEXT, *Mancel R. Ellis* and *D. G. Crawford,* The Ontario Institute for Studies in Education, University of Toronto

The general purpose of this study was to examine the relevance to individualized instruction of the Wallach and Kogan model of "motivational disturbance." The construct 'motivational disturbance' postulated by Kogan and Wallach offers an explanation for the emergence of 'irrational' vs 'rational' patterns of decision-making behavior, associated with dispositions towards risk or conservativism across divergent dimensions of psychological behavior.

The relevance of the Kogan and Wallach model to decision-making in an instructional context has not been made explicit. Therefore, the first objective of this study was to examine its role in a real achievement-related situation in which the incentive aspects of the tasks were perceived by the decision maker to have relevance to an authentic academic achievement-related context. The second major objective of the study was to explore other possible dimensions of the "motivational disturbance" construct.

The subjects were graduate students in education who were asked to complete a number of questionnaires to provide measures of "test anxiety," "defensiveness," "control beliefs," and risk-taking dispositions. A task concerning the relative effectiveness of two instructional techniques in programmed instruction (EGRULE and RULEG) required them to make decisions in a binary choice situation for which they believed correct answers existed. Feedback was provided on the basis of their decisions in fixed schedules of positive and negative confirmation. The extrinsic incentives aspects of the task were defined in terms of points to be gained or lost for correct or incorrect answers, and were determined on the basis of indicated confidence in their decisions.

The dependent variables were derived from performance characteristics which were used as indices of irrational decision-making behavior. Intercorrelations were obtained between the independent variables constituting the construct "motivational disturbance" and the dependent variables.

Implications for the design of individualized instructional environments were drawn.

EFFECTS OF TEACHER TRAINEE APTITUDES ON OBSERVATIONAL LEARNING, *Mary Lou Koran,* University of Texas at Austin, *Frederick J. McDonald,* New York University, *Richard E. Snow,* Stanford University

The present study was conducted to examine the effects of verbal and perceptual dimensions of individual differences in relation to the efficacy of two different kinds of modeling procedures in the acquisition of a teaching skill.

Following the administration of aptitude tests selected from the French Kit of Reference Tests for Cognitive Factors and from a series of specially developed audio-visual testing materials, 121 Stanford intern teachers were randomly assigned to three treatment groups: a film-mediated modeling condition, consisting of a filmed portrayal of Analytic Questioning, the particular teaching skill to be learned; a written-verbal modeling condition consisting of a verbatim text of the sound track from the film model; and a no modeling control group. The criterion performances assessed included the frequency, variety and quality of Analytic Questioning used upon three separate teaching sessions.

Instructional treatment main effects as well as Aptitude x Treatment interactions were investigated. Analysis of variance was used to test instructional treatment main effects. These analyses disclosed highly significant treatment effects in which subjects in both the film-mediated and written-verbal modeling conditions generated a significantly higher frequency, variety and quality of Analytic Questioning than did control group subjects. Similarly, subjects in the filmed modeling condition produced a significantly higher frequency, variety and quality of Analytic Questioning than did subjects in the verbal modeling condition.

Aptitude x Treatment interactions were evaluated by comparing regression slopes for different treatments. Analyses of Aptitude x Treatment interactions indicated that scores on Hidden Patterns, Film Memory and Maze Tracing tests interacted significantly with the instructional treatments. The magnitude of these interactions tended to increase across teaching sessions.

These findings indicate the most effective instructional methods differ from subject to subject with such differences between subjects being correlated with tests of ability. Results such as these, if replicated, may eventually provide decision rules for greater individualization of teacher training programs.

THE RELATIONSHIP OF INDIVIDUAL DIFFERENCES IN THE ORIENTING RESPONSE TO COMPLEX LEARNING IN KINDERGARTENERS, *Mary E. Manske* and *Frank H. Farley*, Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin

The present study investigated the assumption that individual differences in the OR, as measured by heart-rate deceleration, can be used to predict performance in a highly unrelated learning situation, and it further extended previous investigations in the four following ways: (1) to see if the relationship would hold up at the earlier ages of five and six, (2) to see if previously established interactions of OR and sex on learning would be manifest at this early age, (3) to try to establish trends by the use of three OR categories (High, Medium, and Low) in lieu of the previously used High and Low categories, (4) to try and establish the relationship across several learning tasks that are implied to involve different learning processes, namely, concept identification, discrimination learning, and one-trial paired-associate learning.

96 *Ss* were measured on their heart-rate change to a 1000 cps tone. On the basis of his position in the entire distribution of scores, *S* was

classified as either a High, Medium, or Low Orientor. At a later time 96 *S*s received the concept identification and discrimination learning tasks and 65 *S*s received the one-trial paired-associate task.

A quadratic trend was established between the magnitude of the OR and performance on paired-associate learning but the direction of the quadratic trend was highly dependent on sex. Females displayed the relationship in an inverted-U form, i.e., better performance was associated with the Medium OR condition, whereas males displayed the relationship in a V form, i.e., better performance was associated with the High and Low Orientor conditions. Although the shape of the trends was similar for the two other tasks, they were not significant at accepted levels.

Thus, classification of *S*s on the basis of their OR at this early age appears to provide some predictive ability of performance in complex learning tasks when sex of *S* is taken into account.

Session 1.16 **Division E**

CONTRIBUTED PAPERS

Chairman: To Be Announced.

THEORETICAL-PRACTICAL CHARACTERISTICS INVENTORY, TPCI, FACTOR SCORE DIFFERENCES FOR FOUR UNIVERSITY COLLEGES AND NINE ACADEMIC MAJORS, *R. J. Planisek*, Kent State University

The limits in predictive power of ability and achievement measures has stimulated researchers to consider seriously the realm of non-academic predictors. The TPCI is an attitudinal instrument which in previous studies has been shown to have ability, academic performance, and study-habit correlates. The theoretical-practical nature of this instrument and its academic correlates leads one to hypothesize about its ability to discriminate between students of various academic majors. Hence, this investigation is concerned with the discriminatory aspects of the TPCI subscales for different colleges within a university and, specifically, for several academic majors.

The instrument was administered to 599 upperdivison students at Kent State University during the Spring Quarter of 1968. A stratified random sampling was made within the four university colleges as well as within nine academic majors. The majors were chosen for their alleged differences in theoretical and practical curricular contents.

Product moment correlations among the 60 items were factor analyzed by the principal-components method and rotated to the varimax criterion. Four major factors were obtained: Technical-Practical, Scholastic-Theoretical, Application-Practical, and Practical-Collegiate. These accounted for 94.6 per cent of the common variance. Factor scores were computed for each student, from which frequency distributions were obtained for each of the 4 colleges and 9 academic majors. The various group factor score means were compared by the Duncan's New Multiple Range Test. Significant differences among many of the means were found at the .05 and .01 levels. Thus, support was given to the hypothesized discriminatory ability of this instrument.

These findings are encouraging and have provided the rationale for more rigorous studies. Presently the data are being subjected to discriminant and profile analyses. These techniques may provide information which has potential in the counseling and guidance area.

HEADSTART FOLLOW UP 1965-1968: VALIDATION OF AN OBSERVATIONAL INSTRUMENT FOR PREDICTIONS REGARDING SCHOOL SUCCESS, *Ruth Formanek*, Hofstra University

The purpose of the study was to validate an observational instrument for predictions of school success.

I. *1965 Summer.* Behavioral observations by means of specimen description were made daily on 59 preschool children by five observers over an eight-week summer Headstart program. All observations were categorized according to a system proposed by L. Murphy and adapted for classroom situations by R. Spaulding. Categories deal with the child's ability to cope with the demands made on him in free-play and teacher-directed settings, and include the following behaviors: aggression, resistance, independence, productivity, group oriented, passive conformity, and withdrawal. These behaviors were further categorized into "school-adaptive" and "school-maladaptive" types. Peabody and S.B. tests were administered. Predictions of school success, based on the observation of the children's behavior, were made for extreme groups only.

II. *June 1968 Follow-up.* Sample decreased from 59 to 42. In the absence of a more reliable criterion, grade placement was used as an index of school success. All children, except one, were old enough to begin third grade in September 1968. However, only half the sample was found to be at grade level. Two predictions made in 1965 were found to be correct:

1. High amounts of aggressive and resistant behavior correlated negatively with school success; in particular, no child who was found to be highly aggressive in 1965 is on grade level now.
2. Appropriate, success-oriented behavior in school may be defined as passive and conformist when so demanded in teacher-director settings, and independently productive in free-play situations. The six children who showed the highest degree of such differential behavior in 1965 all function on grade level. Predictions for the relation between IQ and differentiated behavior was ambiguous.

   The following behavioral categories did not differentiate those on grade level from those not on grade level:

   a. The degree of overall productive behavior independent of environmental conditions.
   b. The total variability of behavior over the eight-week period of 1965.
   c. The degree of distractible behavior, daydreaming, or fidgeting.
   d. The degree of sociable behavior.
   e. Constant changes over the eight-week period of time in each behavioral trait observed.

Replication of the study is planned on a larger scale using data on aggressive behavior (1) and on differentiated behavior (2) only.

Follow-up will use achievement tests as a criterion in addition to grade placement.

NEED ACHIEVEMENT, FEAR OF FAILURE, PERCEPTION OF OCCUPATIONAL PRESTIGE, AND OCCUPATIONAL ASPIRATION OF ADOLESCENTS OF DIFFERENT SOCIO-ECONOMIC GROUPS, *M. S. Tseng* and *D. L. Thompson*, West Virginia University.

The purpose of this study was to determine if adolescent boys of various socio-economic groups differ significantly in their achievement motivation, fear of failure, occupational aspiration, and perception of the prestige hierarchy of occupations.

A sample of 238 Ss drawn from 9th, 10th, 11th, and 12th grade male students was classified into three socio-economic groups (M,L,LL) based on their father's educational level, mother's educational level, and father's occupation. The instruments used were the Achievement Thematic Apperception Test (TAT), Mandler-Cowan's Test Anxiety Questionnaire (TAQ), the Occupational Prestige Scale (OPS) consisting of 20 occupations adapted from the North-Hatt scale, and Haller's Occupational Aspiration Scale (OAS). The TAT consisted of four pictures (1, 2, 7, 8; Atkinson, 1958) presented in a neutral classroom situation. Scoring was done by two trained graduate students with inter-rater reliability of .90. A short form of TAQ consisted of 32 items. Each item was graded on a 9 point scale with 1 representing low anxiety and 9 representing high anxiety level. The OAS consisted of 8 multiple-choice items designed primarily for use among male high school students. The total score is interpreted as a relative indicator of the prestige level on the occupational hierarchy which an individual views as a goal. The OPS consisted of 20 occupations which were selected from the list of 90 used in the North-Hatt study. The total score is interpreted as an indication of the deviation of the S's perception of occupational prestige in relation to the ideal norm.

ANOVA and Duncan's Multiple Range Tests revealed the following: (1) there was no significant difference among the socio-economic groups on need achievement; (2) group M differed significantly ($p < .01$) on fear of failure from group LL, with group LL showing greater degree of fear of failure; (3) group M differed significantly ($p < .01$) on the perception of occupational prestige from groups L and LL, with group M showing less deviation from the ideal norm; and (4) group M differed significantly ($p < .01$) on occupational aspiration from groups L and LL, with group M showing higher level of occupational aspiration.

CATEGORIZING HIGH SCHOOL GIRLS INTO OCCUPATIONAL PREFERENCE GROUPS ON THE BASIS OF THE DISCRIMINANT FUNCTIONS RESULTING FROM THEIR INVENTORIED INTEREST PATTERNS, *Robert F. Mooney*, Salem State College.

This study demonstrated that significant differences in interest patterns exist among high school college preparatory girls choosing careers assigned to eight broad occupational groups.

1,114 girls were asked to select the occupation they intended to pursue for a life's work. Based on their selection, each girl was categorized into one of eight general occupational groups. These same girls were also given the Kuder E General Interest Survey. This instrument yielded a vector of scores on ten different interests for each girl. Thus, each girl had a ten-score array representing her interest pattern and was categorized into one of eight broad occupational groups.

First, a completely generalized multivariate analysis of variance design determined that, although significant differences in interest patterns existed between occupational groups with high school grade levels partialed out, and between high school grades with groups partialed out, the group differences appeared to be of the same nature across all grade levels.

Then, using two-thirds of the original sample, a multiple discriminant analysis was conducted to determine if significant differences in interest patterns existed among girls in the broad occupational groups. Significant differences were determined to exist. Hotelling's T^2 test was run to determine the nature of the group separations. The discriminant functions were derived in order further to pinpoint the differences between individual occupational groups in terms of interest scores. A "hits-misses" ratio was determined using the other one-third of the sample, resulting in the cross validation of the classification power of the discriminant functions.

The result of this study clearly indicates that, on the basis of their interest patterns, high school girls can be effectively categorized into occupational preference groups more specific than the often reported science-nonscience and college-noncollege preparatory categories.

Session 2.2 **Division C**

PROBLEM SOLVING NO. 1

Lee S. Shulman, Michigan State University, Chairman

ERROR PATTERNS IN PROBLEM SOLVING FORMALIZATION, *John Eliot, George Kennedy,* and *Gilbert Krulee,* Northwestern University

Solution of algebraic word problems requires the student to combine natural language and mathematical symbols. Poor students may differ from able students in their inability to identify the physical assumptions given in the problem statement. Similarly, poor students may differ from able students in their dependence upon information presented early in the problem statement.

The act of solving algebraic word problems can be described in terms of five steps. The two hypotheses of this study were tested by analyzing the verbal and written protocols of 28 students from the 11th grade in terms of these five steps. Students from a "regular" and an "honors" program were asked to solve six different word problems of comparable difficulty. A tape recorder recorded their "thinking aloud" while solving the problems one at a time.

An analysis of the protocols clearly confirmed the hypothesis that poor students differ from able students in their inability to identify the physical assumptions given in the problem statement. Moreover, those students who failed to solve the problems did not begin their solution procedure with information presented early in the problem statement. The second hypothesis was not confirmed.

The results of this study indicate that teachers should be less concerned with getting students to define the relationships between problem elements and more concerned with getting them to identify the physical assumptions made in the problem statement. Further, these results provide some useful guidelines for computer programers interested in designing dialog problems capable of adapting themselves to "converse" with poor algebra students in one way, and with able students in a different way.

A COMPARISON OF THE EFFECTIVENESS OF DISCOVERY VERSUS DIDACTIC METHODS AND TEACHER GUIDED VERSUS INDEPENDENT PROCEDURES IN PRINCIPLE LEARNING, *John T. Flynn* and *Carol Wiesner*, University of Connecticut

The purpose of this study was to determine the relative effectiveness of discovery versus didactic methods for the teaching of spelling principles to elementary school children. The study was also designed to examine the effectiveness of teacher-guided group procedures versus independent procedures used in conjunction with both methods.

Sixteen sixth grade classes were randomly assigned to the four experimental treatment groups. The discovery method with teacher-guided group procedures was used for Group A, the discovery method with independent procedures for Group B, the didactic method with teacher-guided group procedures for Group C, and the didactic method with independent procedures for Group D. Analysis of the students' scores on a standardized spelling test revealed no significant differences in spelling achievement between the groups.

A six week experimental program was developed to teach six basic spelling principles involved in adding word endings. The teaching of the principle differed between groups according to the experimental treatments. The materials and procedures used for reinforcement, testing, and review were identical for all classes.

Two posttests were developed in order to measure spelling achievement relative to the experimental program. One posttest was given at the conclusion of the program, the second after an interval of six weeks. Each test consisted of three parts: Part A was designed to measure retention, Part B, transfer, and Part C, problem-solving. Standard procedures were used to establish the validity and reliability of the tests.

Multiple Classification Analysis of Variance (2 x 2) was used to analyze the results of the posttests for 87 subjects in each of the four groups. No significant differences were found for either main effects or interactions on either the immediate or the delayed posttest. Several reasons were postulated for the results, and suggestions for further research are given.

THE EFFECT OF TWO TYPES OF VERBAL HIERARCHY ON PROBLEM SOLVING, *John T. Guthrie*, Johns Hopkins University

Previous research has failed to account for performance on tasks, such as Maier's 2-string problem, in which Ss are required to combine one or more objects to solve a problem. It is argued that two types of verbal mediators are instrumental in the solution of such problems: uses, which are Ss' verbal responses to the objects necessary for solution, and functions, which are Ss' verbal responses to the verbal statement of the problem. Two experiments, each employing 2^4 factorial designs and including several versions of two unique problems, provided tests of these predictions.

In the first study, each S's verbal mediators to the objects and the statement of the problem were elicited and then his speed of solving two problems was assessed. In the second study an independent group of Ss was required to solve two problems without prior elicitation of the Ss' verbal mediators. In the first study, the dominance of mediators appropriate to solution was ascertained for each S independently; in the second study, the dominance of appropriate mediators was established on the basis of group norms developed from the first study.

In both studies, there were four treatment conditions, which consisted of requiring Ss to solve problems for which their mediators, functions and uses, were either dominant or subdominant in their mediational hierarchies. It was predicted that Ss' speed of solving the problems would be directly proportional to the dominance of the functions and uses in their respective hierarchies. Each experiment demonstrated that problem solving is facilitated both by the dominance of the function ($p < .001$) and by the dominance of the use ($p < .01$). The results confirm the mediational analysis of problem solving.

READING AND LANGUAGE ARTS: ABSTRACT THINKING, A SUB-CAPABILITY OF PROBLEM SOLVING, *Sara W. Lundsteen*, The University of Texas at Austin

Gagné's theory of hierarchial curriculum requires that crucial subcapabilities be located and turned into objectives that can be observed and tested. The goal of the present study was to investigate the cause-effect relationship between trained and untrained variables of abstract quality in children's thinking as related to a learning hierarchy designed to improve creative problem solving. The null hypotheses were: (a) as measured by 3 tests of abstract thinking, there is no significant difference in mean scores for a group receiving 28 lessons designed to develop the abstract quality of thinking at fifth-grade level, and 4 other treatment groups receiving the same amount of special instructional time but with 4 differing programs; (b) there is no transfer or significant difference between the above groups on criterion measures of problem solving, listening, and reading.

A five-group design, with pretests and posttests, was employed. The *Ss*, from a pool of 45 intact classes whose teachers had volunteered for the experiment, were assigned randomly to treatments. The measurement consisted of a battery of 11 tests including 3 experimental tests of qualitative levels of thinking. The findings support the research hypotheses of significant differences between treatment groups on all measures. Analysis of variance, covariance and Scheffé's test were employed. Differences or rankings almost consistently favored the group trained in qualitative levels of thinking when the difference was not in favor of the treatment group especially trained on the transfer variable in question, listening, problem solving or reading. Differences favored the low

I.Q. *Ss* in the qualitative treatment group on the STEP reading total and on 3 out of 4 subscores ($p < .05$). In a gross way the Gagné theory of learning appeared to be supported with respect to the abstract subcapability. For children with low I.Q. (62-91), assistance with the abstract thinking quality may enhance reading ability at the fifth-grade level.

Session 2.3 **Division C**

TEACHER EDUCATION NO. 1

J. Hugh Baird, Brigham Young University, Chairman

IMMEDIATE AND DELAYED FEEDBACK PROCEDURES FOR MODIFYING STUDENT TEACHING BEHAVIOR ACCORDING TO A MODEL OF INSTRUCTION, *Darlene Heinrich,* Beaver College, *Hugh F. McKeegan,* Bucknell University

Frequently supervisor feedback is so delayed and disjointed that it is of little use to the prospective teacher. The purpose of this study was to determine whether, in a situation where a comprehensive model of instruction had already been developed, systematic feedback would improve the application of the model.

Subjects (N-10) were matched according to subject area and were randomly assigned to experimental and control groups. Experimental subjects were given immediate feedback via a series of color-coded cards when instances of the use of (a) advance organizers, (b) defining or describing attributes, (c) inter- and intra-associations, and (d) positive or negative instances were observed. Control subjects were given delayed feedback in conferences conducted immediately after instruction had been completed. It was hypothesized that the discrepancy between what a teacher believes she teaches and what is observed as being taught would be reduced over time (8 weeks) for both groups but that the reduction in discrepancy between belief and observed practice would be greater for the immediate feedback group. It was further hypothesized that student teacher attitude would not be unfavorably affected by the experimental procedure.

A regular observation schedule was set up to permit systematic recording of the efforts of both experimentals and controls to implement the instructional model and to provide immediate feedback to the experimental group. Data analysis was completed via repeated measures, linear trend, and Neuman Keuls' techniques. Attitude responses were tallied and where possible treated via chi^2. Discrepancy scores over successive observations were found to be significant (.05) in the predicted direction for six categories. Positive total and negative total categories indicated a significant linear trend. Discrepancy scores decreased more rapidly for the group receiving immediate feedback and student teachers' attitudes were not unfavorably affected by the experimental treatment.

The study indicated that the behavior of student teachers in applying an instructional model can be affected with a high degree of consistency.

IMPROVING THE COGNITIVE FUNCTIONING OF PUPILS THROUGH TEACHER-TRAINING IN PROCESS OBJECTIVES: A FIELD EXPERIMENT, *Esin Kaya*, Hofstra University

The study purported to determine if teacher-training in incorporating cognitive functions into subject-content curricula would improve the cognitive functioning of pupils.

First, teachers inductively arrived at definitions of the following functions: association, classification, comparison, divergent production, analysis, synthesis, induction, deduction, and evaluation. Second, they formulated behavioral pupil outcomes combining these functions with the subject content they wanted to teach. Third, they modified their instructional techniques to achieve these combined objectives by altering their own classroom behavior.

A field experiment was conducted to determine the effect of teacher training on the pupils. A pre-post-test design with control groups was utilized. Classroom observations of teachers were conducted before and after training to determine the changes in classroom practices. Children in all groups were tested with three main instruments: (a) a battery of selected standard tests of cognitive functioning, (b) validated local achievement tests combining the process objectives with subject-matter, (c) standard achievement tests.

Results confirmed the original hypothesis. (1) Experimental groups achieved significantly better than control groups on tests emphasizing both content and cognitive processes. (2) As was expected, no differences were noted on standard achievement tests. (3) When teachers continued their new instructional technique over one academic year, their pupils' general cognitive ability scores improved significantly; no such improvement was observed when teachers practiced the new technique for the duration of one unit only.

Implications: (1) Children's cognitive functioning may be improved through training. (2) Teacher training readily effects changes in pupils' achievement but must be practiced by the teacher for a year to effect change in pupils' general ability. (3) Standard achievement tests are not relevant to the assessment of cognitive abilities within the context of subject-content.

COMPARISON OF FOUR TEACHER TRAINING PROCEDURES IN ACHIEVING TEACHER AND PUPIL "TRANSLATION" BEHAVIORS IN SECONDARY SCHOOL SOCIAL STUDIES, *Gregg B. Millett*, University of Texas at Austin

The purpose of the study was to determine whether different training procedures could change specific behaviors of intern teachers and their pupils in secondary school social studies classes. Four training procedures were compared in the study: (1) an *unstructured discussion* of material that was to be taught later in a social studies class; (2) *oral instruction* on how to teach the material; (3) a videotaped *demonstration* of how to teach the material; and (4) a combination of the *oral instruction and demonstration* procedures.

The instruction and demonstration training procedures emphasized the development of "translation" of a given piece of material through class discussion. Translation was defined as statements about the meanings of written words and combinations of written words in light of the context in which the words were used. Pupil translation statements, both written and expressed in discussion, were measured. Seven types of teacher behavior were measured: translation, directions, elicitations, acceptances, rejections, probes, restatements, and periods of silence.

The experiment was conducted in the regular program of teacher training in secondary school social studies in the Stanford Secondary Teacher Education Program. Forty-three social studies intern teachers were randomly assigned to the four training groups. Within two days after training all of the intern teachers taught the same material in one of their secondary social studies classes. These discussions were tape-recorded and written translation tests were administered after the discussions.

Blind ratings of the tapes and tests averaged over 80 per cent agreement between two raters. The results indicated significant differences (.01 level) between treatments on the translation behaviors of the teachers and pupils in the discussions. The unstructured discussion procedure was the least effective; the presentation and demonstration procedures were equally effective; the demonstration plus presentation procedure was most effective (.05 level of significance between pairs of treatment means). No differences were indicated between treatments on the written test scores.

The results support the use of demonstrations combined with presentations in transmitting certain complex behaviors to teachers.

Session 2.4 **Division D**

DEVELOPMENT OF MEASUREMENT INSTRUMENTS

Miriam Bryan, Educational Testing Service, Chairman

DEVELOPMENT OF AN INSTRUMENT MEASURING ATTITUDES TOWARD READING IN PRIMARY PUPILS, *Eunice N. Askov*, University of Wisconsin

An attitude inventory was developed to assess change in attitudes toward reading in children in the primary grades. It was designed as part of a larger study examining the effects that a system of individualized reading instruction would have upon achievement and attitudes of primary children.

Prior to the development of the instrument, the favorite leisure-time activities of primary children were determined by interviewing twenty second and third graders of both sexes. Then an artist depicted the most frequently named activities in a boy's and a girl's version. The activities in the two versions correspond to each other as closely as possible.

Pictures of three types of reading activities are each paired with nine pictures of other leisure-time activities, allowing the pupils to choose reading over other activities a maximum of 27 times. Non-reading activities are paired together thirteen times as distracting items. The test is scored by counting the number of times reading is chosen as a preferred activity.

A test-retest reliability coefficient of .90 was obtained with a one-week interim between tests. In a validation study, parents of second and third grade pupils (who had previously taken the inventory) were asked to mark the pictures of the activities that their children, in their estimation, would prefer. The choices of the parents and their children were then compared.

The advantages of such an inventory are perhaps obvious. The children's attitudes toward reading are assessed by a game-like instrument, the purpose of which is not obvious to them. Furthermore, it is easy and quick to administer to primary children since it involves no reading or writing. By making forced choices between pictures of activities preferred by the age group, children reveal their attitudes toward reading as one of a number of other leisure-time activities.

PRESCHOOL ACADEMIC SKILLS TEST, *Esther Kresh*, Pittsburgh Public Schools, *Bert F. Green, Jr.*, Carnegie-Mellon University

The passage of the Economic Opportunity Act of 1964 and the Elementary and Secondary Education Act of 1965 has given great impetus to the development of preschool programs for poverty children.

As part of the Pittsburgh program, the Pittsburgh Public Schools developed the Preschool Academic Skills Test as a diagnostic instrument to facilitate the planning of an appropriate cognitive curriculum for each child. The items for the test were developed on the basis of an analysis of first grade materials to determine those skills that were necessary prerequisites to the skills taught in first grade.

After an initial tryout and revision, the test was administered to 428 four-year-olds from poverty neighborhoods and 326 middleclass four-year-olds from private nursery schools. In addition, the test was administered to a smaller group of 120 five-year-olds.

The following results were obtained:

1. A factor analysis of the interitem phi coefficients revealed that ten factors would account for most of the interitem correlation. These factors were labeled as Vocabulary, Color Naming, Classification, Functional Relationships, Visual Matching, Auditory Matching, Picture Arrangement, Symbol Series, Counting, and Verbal Concepts.
2. Total test reliability measured by Kuder-Richardson Formula 20 was .939. Subtests Kuder-Richardson reliabilities ranged from .488 for a four-item subtest to .842.
3. The multiple correlation of the Preschool Academic Skills Test subtests with the Metropolitan Readiness Test administered a year later was .80 for the Kindergarten sample (N=69).
4. The multiple correlation of the Preschool Academic Skills Test subtests administered at the beginning of Kindergarten with the Metropolitan Achievement Test administered at the end of the first grade was .793 (N=69).
5. Results of the correlation between the Preschool Academic Skills Test administered to the four-year-old group and the Metropolitan Readiness Test will be available by February.

The skills measured by the Preschool Academic Skills Test seem to be strongly related to first grade readiness and first grade achievement and are, therefore, appropriate for consideration in the cognitive curriculum of a preschool intervention program.

THE DEVELOPMENT OF THE ZIP TEST: A QUICK LOCATER INSTRUMENT FOR USE WITH MIGRANT CHILDREN, *Norval C. Scott, Jr.*, University of Southern California; *Beatrice Ward*, Far West Laboratory for Educational Research and Development; *Edwin C. Coffin*, Monterey County Schools

During the 1967-68 school year, a study was undertaken to develop a placement instrument to be used with migrant children.

The instrument was to be designed so that it would (1) assess a child's mastery of a sequence of behaviorally defined reading and mathematics skills so as to determine fairly accurately the learning level (generally, the grade-level book) at which his instruction should begin, (2) assess the English language facility of the pupil, (3) be administrable by any adult (teacher, teacher's aide, or clerk), and (4) take no longer than 15-25 minutes to administer.

The final instrument consisted of three sections: one for English language facility, using stick figures for stimuli; a second for reading, containing a word recognition list, paragraphs for the child to read, and antonym lists; and a third for math with graduated computational problems.

The validity was computed by using the independent judgment of eight teachers as the criterion against which the Zip Test results would be compared. Each teacher was placed in a non-home school environment, and was allowed thirty minutes to interview a migrant child in order to locate his math and reading levels and rate his English language facility. The children to whom the test was administered were selected at random from the class lists of the school. The Zip Test was given to the same children at a later date by a second team of teachers. Altogether, 69 children were involved in the validity procedure.

To determine the extent of departure of the test results from the expectations of chance alone, chi square was used to test for agreement between the findings of the teacher interviewers and the Zip Test in both math and reading placement. For math, $X^2 = 11.36$ ($p < .01$) and for reading, $X^2 = 9.80$ (p. $< .01$). The validity of the English language facility was obtained by a correlation coefficient between the teacher interviewers and the Zip Test scores ($r = .77$, $N = 69$).

Reliability was obtained by a test-retest of the Zip Test, using migrant children in a summer school situation. The correlation coefficients ($N = 53$) were .71 for the English language facility section, .94 for reading, and .97 for math.

The administrations of the Zip Test, thus far, indicate that any competent adult can give the test to a migrant child in 15-20 minutes.

AUDITORY AUTOMOTIVE MECHANICS DIAGNOSTIC ACHIEVEMENT TEST, *Richard A. Swanson*, Bowling Green State University

The purpose of this study was to develop and validate a test to measure the ability to diagnose malfunctions in automobiles through the auditory sense. This test is referred to as the Auditory Automotive Mechanics Diagnostic Achievement Test (AAMDAT).

A list of automobile malfunctions that are commonly diagnosed through the trained ear was compiled. Those malfunctions agreed upon most frequently were built into automobiles. Using a high quality binaural recording-reproducing system, the malfunctioning automobiles were recorded. The recordings, along with recorded instructions, made up the AAMDAT.

Four distinct populations were sampled. The first was the No Experience Group (N = 67). They received the AAMDT only. The second consisted of First Year Auto Mechanics Students (N = 91). They received the AAMDT and an auditory aptitude battery. The third consisted of Graduating Auto Mechanics Students (N = 67). They received the AAMDAT, the auditory aptitude battery, and a paper and pencil auto mechanics achievement test. The last sample consisted of experienced Auto Mechanics (N = 44) who received the AAMDAT only.

The AAMDAT was found to be a reliable instrument for the total group, but not reliable enough to make the fine discriminations necessary within any one of the sub-groups. The concurrent validity of the AAMDAT, which was evaluated by correlating AAMDAT achievement to peer nominations on the same dimension, was partially supported. A principle axes factor analysis on the AAMDAT item produced groups of items having high inter-correlations. Several factors were indentified psychologically. The potential of the Auditory Aptitude Test Battery as a predictor of auditory diagnostic achievement was revealed in the .41 ($p < .0001$) correlation of the nine predictor aptitude tests to the AAMDAT. The independence of AAMDAT and paper and pencil auto mechanics tests achievement was supported.

Session 3.6 **Division C**

COMPUTER ASSISTED INSTRUCTION NO. 1

Norman T. Bell, Michigan State University

LINGUISTIC AND TUTORIAL MODELING FOR NATURAL LANGUAGE CAI, *Fred D. Bennik, Robert M. Schwarz,* and *Harry F. Silberman,* Education Systems, Systems Development Corporation

This research has as its goal the development of an adaptive and generative computer-based tutorial system that can enhance higher order verbal skills through meaningful dialogue, yet relieve a lesson author of much of the burden in lesson preparation that characterizes present CAI systems.

The first phase of research consists of natural language data processing and conceptual modeling designed to support a CAI system. The second phase focuses on developing a model of tutorial decision making that can be embedded in such a system.

Research in the first phase has resulted in the development of a cognitive and linguistic theory expressed as a computational model of verbal understanding. This model has been realized in computer programs that perform functional operations of syntactic-semantic analysis, inferring answers to questions, generation of coherent discourse, and recognition and generation of paraphrases.

Research in the second phase is to find decision rules for generating and sequencing remedial questions or statements. Data are obtained from the verbal messages of a tutor as he monitors and augments the interaction between a student and a computer-administered lesson similar to that in which the natural language CAI system will be engaged. From a study of these verbal protocols, we are seeking a set of effective decision rules for computer generation of sequences of instructional dialogue.

EFFECTS OF COMPUTER BASED INSTRUCTION UPON ACHIEVEMENT IN ELEMENTARY STATISTICS, *James G. Cooper*, University of New Mexico

Concern is growing about computer assisted instruction. The costs of program development remain very high. Conversational Statistics have been offered as a partial solution to the problem. Tutorial instructions are ". . . provided at the user's request. The objective . . . is to assist . . . by suggesting alternative methods of calculation and by providing additional information on statistical methods."

Many students entering their first course in statistics voice apprehension, even despair. It was believed that the experience of communicating directly with a computer and receiving immediate feed-back would result in improved achievement.

It was hypothesized that students exposed to Conversational Statistics via computer terminals will achieve at a higher level than students exposed to a traditional method of teaching elementary statistics.

Four graduate education classes were randomly assigned to one of two treatments. The two experimental groups learned to solve their problems on the IBM 2250, a combination keyboard and scope. The two control groups were instructed in the traditional manner. Three instructors were involved.

A pretest of 8th grade arithmetic problems was given during the first session of each class. Two criteria were employed: the statistics final examination, and an unsigned questionnaire. The latter included such items as:

How do you feel about your statistics class?
How do you feel about your ability to analyze research data?
How do you feel about data processing and research?

Three responses were provided: positive, neutral, and negative.

The data were analyzed with covariance for the pre- and post-test and analysis of variance for the questionnaire.

No significant differences were found for the attitudinal variable.

The data failed to support the research hypothesis. In addition, no significant differences were found for the attitudinal variable.

CHANGES IN YOUNG CHILDREN'S CLASSROOM BEHAVIOR AFTER A YEAR OF COMPUTER ASSISTED INSTRUCTION: AN EXPLORATORY STUDY, *David H. Feldman* and *Pauline S. Sears*, Stanford University

The present study investigated possible effects of a partial treatment of computer assisted instruction on the classroom behavior of first grade children by comparing them with a group who received teacher-led instruction. A Behavior Survey Instrument was developed to permit observation and recording of children's responses to the classroom environment. Analyses of the Behavior Survey Instrument yielded 18 composite scores (9 categories of behavior at two points in time), which were subjected to a factor analysis to test the hypothesis that the behaviors would yield two main factors, an Academic and a Social behavior cluster. Four independent factors (Academic and Social behavior at each of two points in time) were defined. The subsequent data analysis used Ss scores on the four factors as dependent, independent, and correlational variables. Despite independence of the four factors for the whole sample of children (N = 72), subgroup analyses revealed systematic differences between groups.

1. *Stability.* It was found that CAI children were less stable in Academic behavior than non-CAI children, while non-CAI children were slightly less stable than CAI children in Social behavior.
2. *Mean Score Difference.* It was also found that CAI children exhibited a greater amount of Social behavior during the Fall sampling than non-CAI children, while the opposite was true in the Spring. There were no significant differences in Academic behavior behavior between groups in the Fall, while non-CAI children showed a slight decrease between the Fall and Spring behavior.
3. *CAI Reading versus CAI Math.* The results indicated that Fall Social behavior was a reasonably good predictor of the progress a child would make in the CAI Reading Curriculum for the CAI Reading group. This was not as true for the Math CAI children in Math. In the CAI Reading group, a significant correlation was found between Academic behavior and achievement in the *non-CAI subject;* this correlation remained significant even after IQ was partialled out. For the Math CAI children, Academic behavior (Springs) was strongly related to IQ (.71, $p < .01$). Strong positive correlations between behavior and achievement were diminished to nonsignificance when the correlation between IQ and Academic behavior was partialled out. The correlation did hold up in one case, Reading Level, which was a *non-CAI treatment* achievement measure.

The findings were interpreted as suggestive of a possible reduction in the expected positive relations among academic behavior, IQ and achievement in the curriculum in which a child received CAI instruction.

CAI MULTI-MEDIA INSTRUCTION IN PHYSICS: A FINAL EVALUATION OR THE THRESHOLD OF A RESEARCH PROGRAM, *Duncan N. Hansen, Walter Dick,* and *Henry T. Lippert,* Florida State University

Major CAI curriculum research projects are typically justified according to their contribution to 1) technologically-based curriculum developed, 2) evaluative comparisons with conventional approaches to instruction, and 3) contributions to learning and instructional theory. The physics CAI multi-media curriculum, implemented on the IBM 1500 System, is described in terms of the "systems approach" utilized during its development. The rationale and outcomes for media assignment indicate both the contribution of CAI interaction as well as the relative merits of text, film, audio, and CAI presentations. An operational definition of the revision cycle process is provided within the identified learning deficiencies and enhanced operational requirements.

The evaluative comparison with a conventional, collegiate lecture demonstration physics course indicates enhanced terminal achievement levels for the CAI-presented curriculum. In addition, the positive and negative features of interests and attitudes towards this technologically-based form of instruction are reviewed. The evaluative comparison is further reviewed in terms of estimated cost effectiveness for the various instructional presentation forms.

Detailed analysis of the learning protocols are discussed in terms of inferred learning models. The findings on aptitude by media interactions were by and large negligible. The role of anxiety as a determiner of achievement level provides insight into the instructional process. A review of the unresolved questions and suggestive hypotheses for applied learning and instructional theory provides an argument for viewing CAI course development as a vehicle by which to continue a program of research.

Session 3.7 **Division C**

COGNITIVE DEVELOPMENT

Aubrey Roden, State University of New York at Buffalo

VARIABLES INVOLVED IN CHILDREN'S ABILITY TO REPRODUCE SPACE RELATIONS, *Charlotte D. Cox* and *Charles D. Smock*, University of Georgia

A three-dimensional landscape was used to examine a number of variables involved in the child's ability to accurately reproduce space relations. The specific purpose of this study was to examine variation in performance between visual and haptic perceptual modes, static and transformed orientation, and structure of the perceptual field.

Twenty-seven second graders were presented with a modified version of Piaget's landscape task (1956). The task consisted of two identical "countrysides" with Ss required to place objects, one at a time, on their field so as to reproduce the location of each object on the standard field. Each child was administered six trials (randomly selected locations on the field) under static and transformed (subject field rotated 180°) conditions for each the visual and haptic mode. Exact horizontal and frontal coordinates of Ss' placements were recorded.

Magnitude and directional error scores for each independent variable were compared using an analysis of variance. Results indicated no significant difference between the mean error on tasks using visual and haptic perceptual modes. Correlation analysis, however, suggested the two tasks were comparable only under the static condition.

Static and transformed test conditions were significantly different, with transformed being the more difficult. Significant interactions with object locations (items) and with horizontal-vertical, indicate that the effects of relative orientation are modified by the contexts or "field" conditions.

Generally larger mean errors were found on both dimensions, for all items, when the landscape was rotated (i.e., transformed) and the direction of mean error varied with object location.

Only for the static condition was evidence found to support Gibson's and Piaget's comments on the comparability of visual and haptic per-

ceptual modes. The findings in general agree with Piaget's theory of coordination of viewpoints and Pavelko's data related to perceptual support, but these findings suggest that the process of constructing representations of space relations is closely linked to the problem of perceptual or field effects.

CHILDREN'S SPATIAL REPRESENTATIONS AND HORIZONTAL DIRECTIONALITY, *John R. Kershner*, Ontario Institute for Studies in Education and University of Toronto

The purpose of the investigation was to determine the effects of laterality, movement and language on children's ability to store representations of spatial dimensions that are crucial to horizontal directional responding and in so doing: (1) to assess the mode of children's spatial representations, i.e., enactive, iconic, or symbolic, (2) assess the theoretical position of Delacato, Kephart and Piaget, and (3) gain insight into young children's understanding of maps and diagrammatic representations of reality.

A kindergarten class of 24 children was divided into lateralized and non-lateralized groups on the basis of pretest of lateral dominance and the children were pretested for their knowledge of the verbal categories "left-right" both in relation to their own body parts and in relation to their environment. Ss were then randomly exposed to either a participation or spectator treatment condition. In the participation condition the children walked through a 40 foot "T - maze" which was outlined by using wrestling mats on the school gymnasium floor. In the spectator condition the children watched the experimenter walk through the same maze. Posttesting consisted of the children being shown an upright map and an inverted map of the treatment condition and then being tested on their ability to recall and transpose their spatial representation of the treatment condition which involved a horizontal response.

Applying chi square and Mann-Whitney U tests to the data, no significant differences were found across functional knowledge of "left-right" exhibited significantly more successes on both directional tasks. The results were interpreted in support of a language mediation position and tended to question an emphasis upon the influence of motor-perceptual involvement in the development of children's spatial orientation.

THE FACTORIAL STRUCTURE OF REASONING, MORAL JUDGMENT, AND MORAL CONDUCT, *Beth Stephens*, Temple University; *Gene V. Glass*, Laboratory of Educational Research, University of Colorado; *Charles K. Miller* and *John A. McLaughlin*, Temple University

To permit study of the development of reasoning, moral judgment, and moral conduct in retardates and normals, a battery of measures was selected which included 21 of Piaget's and Inhelder's measures of reasoning, seven measures of moral judgment adapted from Piaget's work, and six measures of moral conduct developed by the investigator. Scoring systems were devised for these measures; inter-rater reliability ranged from .81 to 1.00. The tasks were then administered to 75 normal (IQ 90-100) and 75 retarded (IQ 50-75) subjects (total N = 150) enrolled in public schools in the Philadelphia area. Ages for the randomly selected sample were:

25 mentally retarded subjects, ages 6 to 10 (12 male, 13 female)
25 normal subjects, ages 6 to 10 (12 male, 13 female)

25 mentally retarded subjects, ages 10 to 14 (13 male, 12 female)
25 normal subjects, ages 10 to 14 (13 male, 12 female)

25 mentally retarded subjects, ages 14 to 18 (12 male, 13 female)
25 normal subjects, ages 14 to 18 (12 male, 13 female)

Scores were also obtained on the 12 sub-tests of the Wechsler Intelligence Scale for Children, total IQ, mental age, chronological age, the Wide Range Achievement Test, and Warner's Index of Social Characteristics. Intercorrelations of these variables were subject to factor analysis separately for (1) the total group, (2) normals, and (3) retardates. Initial common-factor solutions were rotated analytically to oblique simple structure by the isopromax method. Factors thus derived are interpreted and compared across samples of subjects. The study appears unique in its attempt to factor analyze Piagetian assessments and to determine the relationships which exist among measures of reasoning, moral judgment, and moral conduct. Discussion centers on results obtained from the factor analysis.

Session 3.8 **Division C**

PROBLEM SOLVING NO. 2

Thomas J. Shuell, State University of New York, Buffalo

THE EFFECTS OF ATTENTION ON THE LEARNING HIERARCHICALLY STRUCTURED MATERIAL, *J. Maurice Mahan*, University of Illinois; *J. William More*, Bucknell University

It has been felt for some time that in order for learning to take place within a conceptual hierarchy it is necessary that material be carefully sequenced (Skinner, 1958; Gagne, 1962; and Ausubel, 1965). Recently, however, data have been presented which seem to indicate that this careful sequencing may not be necessary (Roe, Case, and Roe, 1966; Wodtke, et al, 1967; Hamilton, 1963; and Levin and Baker, 1964).

The purpose of the present study was to test a theory concerning the effects of reinforcement in an instructional sequence on attention during learning.

A program on mathematical sets was given to 90 six grade students. The Ss were given the program in one of three forms (ordered, scrambled, or reverse sequence).

The results indicated that there were no statistically significant differences between the ordered and scrambled groups in the number of errors during acquisition or in the time to complete the program. There were significant differences in achievement between the means of the reverse sequence and those of the other two sequences. There were also significant differences in total reading time between the means of the reverse sequence and those of the other two sequences.

It was concluded from this study that the reinforcing properties of instructional sequences affect the amount of attention which a learner gives to the material during acquisition. The results also indicated that the reinforcement in an instructional sequence is not limited totally to knowledge of results, but may be an interaction with another factor such as meaningfulness. The results of this study and of earlier scrambling studies indicate that more effective instructional sequences could be developed by combing a highly structured order with a means of producing a high level of attention.

KINDERGARTEN CHILDREN'S USE OF VOCAL LABELING RESPONSES, *Robert H. McNeany* and *Even R. Keislar*, University of California, Los Angeles

The purpose of this study was to assess the relative effects of instructing kindergarten children to speak out loud relevant or non-relevant labels during a problem solving task involving selective learning, in comparison with children not instructed to label.

The study was carried out with 42, five-year-old Negro children from lower socio-economic class homes. The children were randomly divided into a Relevant-Labeling Group, a Non-Labeling Group, and a Non-relevant-Labeling Group each containing 14 subjects.

For each problem, the child was shown three pictures of an object identical in all respects except size, thickness, or length (e.g. a large cow, a medium-size cow, and a small cow). The child was told which picture was correct; he then had to identify the correct one one each of three additional cards showing the same three pictures randomly rearranged. Children were trained and tested individually over a four day period, 15 minutes a day. During training, children in the Relevant-Labeling Group were instructed to supply the appropriate label (i.e. *Big, Medium,* or *Little*) for each picture they selected. The Non-relevant-Labeling Group was instructed to label the object shown (i.e. *Cow*). The Non-labeling Group was given no instructions to respond aloud. On the posttest, no instructions regarding speaking were given.

Analysis of variance showed significant differences among the three groups both during training and on a posttest involving similar but new problems. The Relevant-Labeling Group was significantly superior to the Non-relevant-Labeling Group during training and on the posttest, but was significantly superior to the Non-labeling Group only during training. It was concluded that instructing children to say aloud relevant instead of non-relevant labels enhances their ability to deal with new problems.

YOUNG CHILDREN'S USE OF LANGUAGE IN INFERENTIAL BEHAVIOR, *Carolyn Stern* and *Evan R. Keislar*, University of California, Los Angeles

Children from impoverished environments are usually handicapped in the mediational use of language. The present experiment was designed to determine whether children from poor families could learn to use sentential connectives and quantifiers in problem-solving tasks involving simple logical operations. A related question was whether they could learn to process this type of information so as to be able to tell whether or not they had been given sufficient bases for drawing inferences. Two methods of instruction were tested.

120 children from 1st and 2nd term kindergartens, blocked on 2 levels of mental age, were randomly assigned to 1 of 3 treatments: (1) a non-oral

group listened to taped commentary and responded by selecting the appropriate 1 of 3 pictures; (2) an oral group listened to the same commentary but was required to speak aloud the relevant verbal cues before selecting the picture; and (3) an uninstructed control. Over an 11 week period, groups of 5 or 6 children were given daily programmed lessons, each lasting approximately 12 minutes. Booklets with special feedback ink and, occasionally, demonstrations and games requiring manipulanda, were used.

Phase 1 of the program taught children to say "can't tell" in situations where there was insufficient information to warrant a "yes" or "no" response. Phase 2 was a sequence of lessons about the octopus; and Phase 3 involved the use of conceptual rules using sentential connectives: *not, and,* and *or*. Pre- and posttests were individually administered to all children.

Results indicated that the instructed groups were reliably superior ($p < .001$) to the controls in the use of quantifiers and connectives. However, teaching children not to guess (i.e., to say, "can't tell" when given insufficient information) was a very difficult task, under either treatment condition. In contrast, the oral group was reliably superior to the non-oral in learning the nature study concepts.

TEACHER DECISION MAKING: THE INFLUENCE OF AUTHORITARIANISM AND TIME ON REASONING, *Elwood B. Traylor*, Wichita State University

This study examined the decision making process of prospective teachers to determine the extent time and authoritarianism influence the decisions about the validity of categorical syllogisms with attitudinally relevant conclusions. Conceptually, these decisions are viewed as the result of intellectual skills and attitudes about the conclusion. Research has (Traylor, 1968) indicated that prospective teachers could improve their ability to make these decisions by strengthening intellectual skills and the induction of a set. The questions asked are: does carefully guided practice in solving syllogisms and in the induction of a set improve reasoning about attitude structured syllogisms for authoritarians and does this improvement persist over a period of time?

The hypotheses may be summarized as follows: (1) carefully guided practice in the solution of syllogisms and the induction of a set will improve reasoning about attitude structured syllogisms by low more than by high authoritarians; (2) the improvement will reduce with time.

100 Ss from 3 classes were used. The MTAI and California F Scale were administered. Selected statements from the MTAI were used as conclusions for syllogisms constructed for the Syllogian Test, which was administered mid-semester to determine attitude errors. Logically valid conclusions that S disagreed with and judged to be invalid and logically invalid conclusions that S agreed with and judged to be valid were classified as attitude errors. Authoritarianism was determined by the F Scale. High and low authoritarians were assigned at random to either immediate or delayed groups. Programed instruction was given each S; the Syllogism Test and set was given immediately after the program or delayed one week.

The data were analyzed to determine the influence of authoritarianism and time on the reduction of attitude errors. Analysis of covariance of a 2 X 2 table was used to test the hypotheses. Both hypotheses were supported.

Session 3.13 **Division D**

FACTOR ANALYSIS—A TECHNIQUE AND A TOOL

Warren G. Findley, University of Georgia, Chairman

USE OF SEQUENTIAL FACTOR ANALYSES TO CLARIFY INTERPRETATION OF UNDERLYING RELATIONS, *Warren G. Findley*, University of Georgia

Definitive interpretation of a factor analysis is questioned by some because the constructs proposed by the investigator often appear arbitrary. An effort to meet this criticism and reduce one element of arbitrariness has been to adopt rules for extracting factors, such as use of the method of principal axes for determining unrotated factor loadings, varimax rotation for maximizing zero entries to produce simple structure, and an eigenvalue criterion for how many factors to rotate.

It is the thesis of this paper that the availability of efficient computer programs makes it advisable to try not merely the number of factors dictated by eigenvalues, but successively additional factors until the meaning of the pattern is clarified.

One problem is presented in which successively 3, 4, 5, and 6 factors were extracted from a table of intercorrelations of grades at West Point in 14 subjects. The rotated factor loadings established a meaningful pattern with 4 factors, one more than dictated by eigenvalues. The first two factors, verbal subjects and drafting subjects, emerged when only 3 factors were extracted, but the third factor was complex; the third factor broke up neatly into two factors, mathematical subjects and subjects studied as upperclassmen, when 4 factors were extracted; additional factors proved singular.

A second problem is presented in which 2 and then 3 factors were extracted and rotated to interpret the intercorrelations of responses to 12 questionnaire items based on official statements of APGA on human rights. The first two factors, counselor support of counselee development and responsibility to the public, emerged when only 2 factors were extracted; an additional third factor, showing special similarity of responses to two items, emerged when 3 factors were extracted, clarifying the significance of the first factor and justifying limitation to the two factors dictated by the eigenvalues.

Principles governing interpretation are offered.

AN ILLUSTRATION OF A FACTOR ANALYTIC INTERPRETATION STRATEGY, *Margaret Harris*, University of Wisconsin

The purpose of this paper is to illustrate the use of a strategy suggested by Chester W. Harris (1967) for determining the common factors in a set of data. He suggested using several different computing algorithms for the initial solution, obtaining both orthogonal and oblique derived solutions, comparing the results, and regarding as the important substantive findings those factors that are robust with respect to method. This paper illustrates a way of comparing the results.

The factor results used for this illustration are the reanalyses, by seven different methods, of the data of two of the Guilford studies as reported by C. Harris (1967).

The procedure involves attempting to find the factors that are the same or similar over solutions by starting with a factor from one solution and matching a factor from each of the other solutions with this, if possible. Those variables with coefficients greater than .30 (absolute) were considered. A comparable common factor was defined as one for which two or more of the same variables have coefficients greater than .30 on that factor for at least five of the seven solutions. In this way, no single initial method could cause a variable to be rejected, and no single rotation method could account for a variable's acceptance.

The reanalyses of two of the Guilford studies will be used as illustrations. For one of these, there is fairly close agreement among the various factor solutions. For the other, the various solutions agree in part, but for some of the factors the results are quite diverse.

A THREE-MODE FACTOR ANALYSIS OF STUDENTS' PERCEPTIONS OF A UNIVERSITY, *J. Christopher Reid,* University of Missouri

The purpose of this study was to analyze the structure of students' responses to certain facets of the university they were attending. A semantic differential instrument having 14 concepts (Library, Dating, Graduate Students as Instructors, etc.) and eight scales (authoritarian—permissive, childish—adult, etc.) was administered to 806 freshmen and sophomores.

The semantic differential instrument was analyzed by Ledyard Tucker's three-mode factor analytic procedure (in *Psychometrika,* September 1966), using "Method III." Estimates of communalities were obtained by a method similar to Thurstone's highest off-diagonal element. The resulting R, P, and Q matrices were rotated to a varimax criterion, and the inverse transformation matrices were used to determine the transformed core matrix. Factors were identified and named for the R, P, Q and core matrices. It was concluded that Tucker's three-mode procedure is not only applicable to semantic differential data, but that the additional information gained for interpretation of responses makes its use worthwhile.

COMPARATIVE FACTOR ANALYSES OF SEVERAL INTERPROFILE SIMILARITY INDICES, *Wilson H. Guertin,* University of Florida

In 1962, Nunnally presented a factor analysis of raw score cross-products. His contribution lies in illustrating the relationship between the derived factor matrix and the matrix of $\sqrt{D^2}$ or D's that was factored. However, his presentation is misleading, because he did not go back to raw score profiles. Actual factor analysis of the matrix of D's was not attempted.

The present analysis of a set of hypothetical profiles clarifies the above relationships further. Five matrices of interprofile similarity indices were factor analyzed. Results are presented in the form of Varimax rotated factor matrices for comparisons.

All interprofile matrices seem to be Gramian and therefore suitable for factoring. Estimation of the diagonal values in the interprofile matrix is discussed. Raw score cross-products factor analysis results were clearly different from the other similarity analyses. The factor matrix from D's gave a clearer separation of profiles than did the factor matrix from D^2's. Cattell's coefficient of pattern similarity showed no advantage over D.

Session 3.14 **Division F**

INSTITUTIONAL FOUNDERS

Merle L. Borrowman, University of Wisconsin, Chairman

BOOKER T. WASHINGTON: PILGRIM'S PROGRESS AND THE PERILS OF PIECEMEAL REFORM, *Jonathan C. Messerli,* Hofstra University

He was born a slave and rose to become the most famous black man of his time. In order to accomplish this remarkable ascent, Booker T. Washington devised his own educational ladder. It began with the crudest self-taught reading lessons, continued with middle class Yankee amenities under the prim and ever alert supervision of Mrs. Viola Ruffner, and culminated in vocational training under the inspirational teaching of General Samuel Armstrong at the Hampton Institute. Absolutely certain of his own means of success, Washington was equally certain that it would work for others of his race. The school which he built at Tuskegee, Alabama, would institutionalize his ideas, and through its graduates he would spread his gospel of "self-help" throughout the South.

For Washington, social reform was education writ large. Rarely radical in his thinking and never a militant, he charted an educational course for his people which began with a training for basic literacy and vocational skills which would lead hopefully to complete legal, economic, and social equality. Succinctly stated, at the outset he believed it was more important for the ex-slave to have the right and skill to earn a dollar than the right to spend it in an opera house. Thus he embraced a reform philosophy of gradualism which was based upon an optimistic vision of the United States as a land of unbounded opportunity for both white and black.

To restudy Washington's efforts is not to mount yet another attack upon him as an educational "Uncle Tom," but to analyze the social milieu of post Civil War America and recognize the ways in which it both encouraged his piecemeal reforms and limited their effects.

BOOKER T. WASHINGTON: RESPONSE, *Clarence J. Karier,* University of Rochester

DANIEL C. GILMAN: FIRST PRESIDENT OF THE CARNEGIE INSTITUTION OF AMERICA, *David Madsen,* University of Washington

On the eve of his retirement as the first president of Johns Hopkins

University, there came to Daniel Coit Gilman the opportunity to direct the fortunes of the Carnegie Institution of Washington, an enterprise that promised to become the most important research institution in the United States. How influential was Gilman in drafting the purposes, in defining the organizational structure, and in determining the research priorities of the new venture?

Answers to these and other questions were sought among a variety of documentary sources, including the papers of Andrew Carnegie at the Library of Congress, those of Daniel Gilman at Johns Hopkins, the Andrew D. White papers at Cornell University, and the John S. Billings papers at the New York Public Library. Of most value, however, were the records, letters, trustees' minutes, and reports of the Carnegie Institution of Washington.

Gilman's resignation in 1904 was the result, in part, of his frustration at his inability to influence the affairs of the institution. Although he was one of those present at the formulation of the purposes and organizational structure of the new venture, he assumed that the role of the president and the relationship between the president and the institution's executive committee were to be approximately the same as that found in the influential private universities. He believed, therefore, that the president was to "consider and present to the Board [of trustees], or to a committee of the Board, for consideration and approval, his matured recommendations . . ." and then to see to the execution of all decisions. However, the members of the executive committee, led by the chairman of the board, John S. Billings, and supported by the institution's secretary, Charles D. Walcott, and Trustee S. Weir Mitchell, understood the institution's by-laws to give charge of institutional affairs to the executive committee, with the president responsible for seeing that the decisions of the committee and of the board of trustees were executed. As a result, Gilman's influence on the course of institutional affairs was sharply reduced and that of Billings and Walcott enhanced. Research priorities were set, and even the day-to-day operations of the institution were overseen by both Billings and Walcott, as well as Gilman.

In addition, the study included examination of Carnegie's motives and of the attitudes of the scholarly community toward the new enterprise.

DANIEL C. GILMAN: RESPONSE, *Mark Beach*, University of Rochester

Session 4.3 **Division C**

COMPUTER ASSISTED INSTRUCTION NO. 2

David D. Starks, University of Michigan, Chairman

PROBLEM-SOLVING ON A COMPUTER-BASED TELETYPE, *Elizabeth F. Loftus*, *Patrick Suppes* and *Max Jerman*, Stanford University

The purpose of the present study was to examine the factors involved in teaching fifth-grade students how to use the computer as a desk calculator to solve mathematical word problems. The program assumed a basic understanding of the four arithmetical operations: addition, subtraction, multiplication, and division.

The 27 students used in this study were taken from an accelerated mathematics group composed of bright fifth-grade students from four different elementary schools. The students had received teletype instruction previously, so familiarizing them with the machine was not a problem.

Sixty-eight mathematics word problems were prepared. To solve these problems the students used the rules they had learned in an instruction set. It should be noted that the students were not allowed to use pencil or paper when working on the teletype. Each exercise was worked on the machine so that all responses could be recorded. After the student worked a problem and indicated his final answer, the computer evaluated the answer. If the answer was correct, the computer typed the message "correct," and went on to the next problem. If the final answer was incorrect, the computer typed "answer is wrong" and went to the next problem. The students worked for six minutes per day until they had completed the entire set of problems.

In addition to the teaching function, the study was designed to permit investigation of variables that might contribute to problem difficulty. A linear regression analysis of the set of problems revealed that most important was whether or not a problem was of the same type as the problem that preceded it. Other variables investigated were number of words in each problem, the minimum number of steps required for solution, and the minimum number of different operations required.

The predictive results of this first analysis are somewhat promising. There is considerable difficulty in intuitively rank-ordering the proportions of correct responses obtained in word problems. We believe that these results give a sense of the real possibility of analyzing and predicting, in terms of meaningful variables, the response performance of children who are solving arithmetical word problems. At first glance, the problem set appears to be quite complex. Yet, with a few variables we have brought a considerable amount of order to it.

THE FEASIBILITY OF COMPUTER ASSISTED INSTRUCTION IN US ARMY BASIC ELECTRONICS TRAINING, *Alexander A. Longo* and *Vincent P. Cieri*, US Army Signal Center and School, Fort Monmouth, New Jersey

This study investigated the feasibility of using Computer Assisted Instruction (CAI) as an instructional method to teach Basic Electronics.

A course covering 2 days of instruction totaling 11 hrs. and 15 min. was programmed for implementation on the IBM 1500 Instructional System using the Coursewriter II Language. The evidence for the effectiveness of this implementation was based on empirical data collected under a controlled experiment which compared CAI with 2 types of conventional training methods currently in use. The sample was composed of 54 students drawn from 3 incoming classes to Basic Electronics Training totaling approximately 150 subjects. The sample was subdivided into 3 aptitude levels (high, medium, and low) and 3 methods of training (CAI, instructor controlled, and TV controlled). This yielded a 3 x 3 complex ANOVA Design.

The results demonstrated that, based on an 85 item criterion measure, there was no significant difference between instructional methods nor was the interaction significant between aptitude level and instructional methods. However, the time to complete this course is related to aptitude level. Using CAI, the average time for the high aptitude subjects to complete with 11¼ hours of instruction was 5 hrs. and 15 min.; the middle aptitude group averaged 9 hrs. and 22 min.; and, the low aptitude group averaged 15 hrs. A weighted average time to complete the CAI course was 9 hrs. and 58 min., for a 11.5% time savings in comparison with 11 hrs. and 15 min. for conventional instruction. These findings demonstrated that CAI can teach basic electronics as well as conventional instruction, and in less time.

A STUDY OF COMPUTER-ASSISTED, MULTI-MEDIA INSTRUCTION AUGMENTED BY RECITATION SESSIONS, *Kenneth Majer* and *Duncan Hansen*, The Florida State University, Tallahassee

Most evaluators of CAI or multi-media instruction have been primarily concerned with establishing the effectiveness of new instructional techniques in comparison with traditional classroom-lecture presentations. Additional questions are examined in this study: If CAI were augmented by face-to-face instructional sessions, what would be the nature of the interaction between professor and students? Additionally, who would benefit more from these extra sessions, the unusually bright or the abnormally slow student?

The second part of this study is an attempt to establish a more efficient use of both conventional and computer-assisted, multi-media instruction by matching the student with the instructional mode best suited to him. This selection procedure is being established through the use of multiple regression and discriminant function analyses. Independent variables are the results of the Attitude Inventory, the ETS College Student Questionnaire, the FACT Inventory and aptitude scores.

Finally, this study examines and compares the performance of students who receive conventional instruction against those who receive computer-assisted, multi-media instruction augmented by recitation sessions.

ERRORS AND LATENCY OF RESPONSE IN COMPUTER-ASSISTED LEARNING AS A FUNCTION OF ANXIETY AND TASK DIFFICULTY, *Harold F. O'Neil, Jr., Duncan N. Hansen,* and *Charles D. Spielberger*, Florida State University

Performance on a Computer-Assisted learning task and changes in state anxiety (A-State) were investigated for 58 college males with extreme scores on the trait anxiety (A-Trait) scale of the State-Trait Anxiety Inventory. An IBM 1500 system presented the learning materials and the A-State scales and also recorded responses and response latencies. CAI materials developed by O'Neil, Spielberger and Hansen (1968) were employed in this study, but with a counterbalanced design; i.e., difficult materials followed easy materials (D-E), or vice versa (E-D).

Results confirmed the findings of O'Neil et al. (1968) in that: (1) A-State scores increased while students worked on difficult materials and decreased when they responded to easy materials; (2) High A-Trait (HA) *S*s responded throughout the task with higher levels of A-State than Low A-Trait (LA) *S*s; and (3) high-State *S*s made significantly more errors on the difficult materials than low A-State *S*s. In addition, a trait by order interaction was found; HA *Ss* had higher A-State scores in the D-E order than in the E-D order, whereas order had little effect on A-State for LA *Ss*.

In general, longer latencies were found on the difficult task, and low A-State *Ss* responded more slowly than high A-State *Ss* when making errors. Also, in the D-E order, low A-State *Ss* had longer latencies for errors for the difficult materials than did high A-State *Ss*. In the E-D order, there was no significant difference in response latencies for the low and high A-State *Ss*.

These results demonstrate that high A-State *Ss* made more errors and responded more impulsively than low A-State *Ss*, and that low A-State *Ss* appeared to adopt the better strategy of spending more time to process the information provided. These findings are discussed in terms of Spielberger's State-Trait Anxiety Theory.

Session 4.5 **Division C**

INDIVIDUAL DIFFERENCES NO. 1

Frank H. Farley, University of Wisconsin, Chairman

INDIVIDUAL DIFFERENCES IN REMINISCENCE AND PERFORMANCE, *Frank H. Farley*, University of Wisconsin

An important concern in analyses of perceptual-motor performance and reminiscence has been the attempt to identify sources of individual differences variance. It has been hypothesized that extroverts and introverts differ in the rate of accumulation and dissipation of reactive inhibition, and accordingly should differ in perceptual-motor parameters, such as reminiscence, that are held to reflect the influence of inhibition.

It has not been possible to determine from previous studies if the greatest contribution to the reminiscence difference has been depressed pre-rest performance or elevated post-rest performance on the part of the extroverts. It has been argued that introverts develop greater conditioned inhibition during involuntary rest pauses in pre-rest massed practice, and from a consolidation hypothesis of reminiscence it would be expected that over a programmed rest interval the conditioned inhibition (resting habit) would consolidate and render the post-rest performance of the introverts significantly inferior, at least initially, to that of the extroverts, with no pre-rest differences being expected.

Six hundred and twenty-three male technical college *Ss* were administered the Eysenck Personality Inventory, from which were drawn 90 introverts and 110 extroverts. These were then tested on pursuitrotor tracking under conditions of 5 minutes massed practice, 10 minutes rest, 15 minutes massed practice.

The two groups were compared on the last trial pre-rest (N.S.) and the first trial post-rest ($p < .05$). Reminiscence was found to be significantly greater in extroverts ($p < .03$ one-tailed test). Over-all pre-rest curves did not differ. Post-rest curves were compared between groups in three 5 minute sections. Groups differed on the linear component in the first section but were undifferentiated in the remaining two sections. In none of the analyses did E-I account for much variance.

The data supported the conclusion that consolidation of the greater conditioned inhibition in introverts accounted for the reminiscence and post-rest differences.

It was concluded that extroversion-introversion is related to perceptual-motor performance and reminiscence, but that the magnitude of this relationship is small.

WAS EBBINGHAUS AN EXTROVERT?, *John P. Gaa* and *Frank H. Farley*, University of Wisconsin

The present research was concerned with developing an account of memory along individual difference (ID) lines. It was hypothesized that a relationship existed between the personality dimension of extroversion-introversion (E-I) and level of activation or arousal. On the basis of previous work showing that high arousal during learning impaired short-term retention but facilitated long-term retention, whereas low-arousal learning led to good short-term retention but marked classical forgetting over longer time periods, it was further hypothesized that E-I would be related to retention. Introverts were held to be generally characterized by higher levels of CNS activation than were extroverts. This assumption is supported by independent evidence. Accordingly, it was predicted that introverts would demonstrate inferior short-term retention but superior long-term retention relative to the performance of extroverts.

Two studies were conducted. Both employed undergraduate students defined as introverted or extroverted on the basis of extreme scores on the Eysenck Personality Inventory Extroversion scale. The first study was a 2x2 design, with unequal Ns due to subject attrition. Immediate and 48 hour retention intervals were used, with 9 extroverts and 7 introverts in the former test and 5 extroverts and 7 introverts in the latter test. Two paired-associate (P-A) (word-number pairs) trials constituted original learning (OL). The results were in the predicted direction, but not significant at conventional levels.

The second study replicated the first on a larger sample, with improved procedural features and greater control over rehearsal through a change in the filler task following OL. 30 introverts and 30 extroverts were tested, in a 2x2 design with 15 *S*s per cell, on immediate and 24 hour retention. The predicted interaction of personality and retention interval was found, but it only approached conventional levels of significance.

Unpublished studies recently undertaken by other investigators support the predicted interaction effect.

The data are discussed in relation to subject selection factors in verbal retention studies, including the topic of the present title, and in regard to the study of IDs in learning and retention.

AUTOMATIC COLLEGE TUTORING FROM STUDENT-FACULTY LINGUISTIC DESCRIPTOR PROFILE SIMILARITY, *Herbert Garber*, S.U.N.Y. College at Oswego

This study attempted to discover if student understanding of verbally-presented college level course content could be improved through advance technical term understanding.

Professors and students in a multi-section course in educational psychology supplied restricted verbal associates, or descriptors (D), to a list of technical terms (TT). Non-overlapping D's from professor and student lists of associates were presented in a separate rating scale for each of 20 TT. Students and faculty rated the relevance of each D to its TT. A computer calculated the standardized professor ratings (SPR) for each D, compared each student raw rating to the corresponding SPR, and, where differences were found, computed a difference score (DS) which varied in size as a function of inter-professor agreement and student disagreement with the professors.

Student total DS's were correlated in three separate series with raw score grades on a regular course quiz which covered, in part, concepts whose TT's were included in the rating scales. The series of correlation coefficients were generated by systematically increasing weights applied to DS's in equal steps over three increment ranges.

The highest correlation found from the three series of generated correlations was —.355 ($p < .01$). This coefficient appeared when a multiplier weight of 2.5 was used to adjust DS's which differed from professor ratings by 50% of the scale range for professor ratings and where high inter-professor agreement existed.

These results indicate that a technique has been developed which may enable individualized instruction to be based on specifiable semantic components of classroom and textbook verbal content. Basic research in teacher explaining behavior effectiveness and in the teaching of verbal concepts, principles, and problem-solving skills generally, may be aided by such an objective measurement device. Also, auto-instructional media can easily evaluate student responses to semantic as well as substantive stimuli and branch to a vocabulary teaching routine if student errors have been diagnosed as "lexical inadequacy" for specific technical terms in the main program.

A REPLICATION AND EXTENSION OF TEPLOVIAN PSYCHOLOGY AND ITS IMPLICATIONS FOR THE STUDY OF INDIVIDUAL DIFFERENCES IN LEARNING, *Herbert H. Severson* and *Frank H. Farley,* Wisconsin Research and Development Center for Cognitive Learning

The principal Soviet approach to the experimental analysis of individual differences (IDs) has been that of B. M. Teplov and his associates. Teplov has identified ID dimensions of "strength," "balance" and "mobility" held to reflect properties of excitation-inhibition in CNS functioning. The present study represents the first extensive replication and extension of Teplovian psychology outside the Soviet Union. It is the first stage of a program designed to integrate this approach to IDs into a full-scale attack on IDs in human learning.

This first study, using graduate students in educational psychology, employed factor analysis in an investigation of the "strength" dimension using tasks most of which were chosen on the basis of a high loading on a "strength" factor from a study by V. D. Neblitsyn. The measures included absolute visual threshold, visual threshold using the "exhaustion method," visual threshold with an auditory distractor, absolute auditory threshold and auditory threshold with a visual distractor. One of the principle measures of strength was the "shape of the curve" index. An additional task was reaction-time as a function of luminance of the visual stimulus. To determine whether strength was related to Western measures

of arousal, the threshold of fusion of paired-light flashes, a relatively well validated index of CNS activation, was ascertained.

A separate study of the retest reliability of the strength measures was also undertaken, with results indicating generally high stability estimates.

The "strength of the nervous system" variable described by Teplov is the primary dimension being hypothesized in this study. With this dimension established, the remaining dimensions will be similarly studied, and the implications of these dimensions for human learning, as discussed in the present paper, will be pursued.

Session 4.6 **Division D**

EMPIRICAL RESEARCH ON OBJECTIVE AND SUBJECTIVE TESTING

Leonard Cahen, Educational Testing Service, Chairman

MEASURES OF RISK TAKING ON OBJECTIVE EXAMINATION,

Malcolm Slakter, Roger Koehler, Sandra Hampton, State University of New York at Albany; *Robert L. Grennell,* State University College at Fredonia

Although risk taking is becoming an increasingly interesting variable in educational and psychological research, very little is known about its relation to age or sex. The purpose of this study was (a) to devise measures of risk taking on objective examinations (RTOOE) that would be suitable for use in grades 5 through 11, and (b) to administer the measures to suitable grades in order to observe the relation of RTOOE with grade level and sex.

The risk measures were based upon the use of nonsense items, where a nonsense item is defined as one that has no correct (or best) answer and no incorrect answer for the given population. Previous research had indicated that five nonsense items embedded in five legitimate items would provide suitable test characteristics. In addition, since there is also evidence that RTOOE is a general trait (i.e., across different types of examinations), convenient synonym-antonym vocabulary items were the type utilized in the measures. The following is an example of a nonsense item used in the measures (Ss were directed to indicate whether the words had the same or opposite meaning, and were informed of the penalty for guessing):

7. marnel.........................mild

In order that grade trends could be examined, the nonsense items which formed the basis of the risk measures were constructed so that they could be used at all grade levels. The legitimate items were varied to be suitable to the various grade levels. The risk score assigned to a S was the proportion of nonsense items attempted.

Split-half reliabilities (odd versus even, adjusted by the Spearman-Brown Formula) for the risk measures were comparatively high across all grade levels, confirming findings from other studies. A sex by grade analysis of variance on the risk scores was performed in order to examine the relations of RTOOE with sex, grade, and the sex by grade interaction.

NEGATIVE STEPS IN MULTIPLE-CHOICE TEST ITEMS, *Carmelo Terranova*, State University College, New York

The constantly increasing dependence upon multiple-choice tests as bases for decisions requires that the rules for constructing such items be grounded in empirical findings. This study attempted to provide some experimental evidence regarding the use of negative stems in two-alternative multiple-choice test items.

A factorial design was used which incorporated (1) the same items stated either positively or negatively; (2) the frequency of change of direction-set (FCDS), which refers to the number of times *S* had to change from considering positive items to considering negative items, and vice-versa —0, 1, 2 changes were examined; and (3) four grade levels (5, 7, 9, and 11th grades). The score obtained and the amount of time needed to complete the experimental instrument (three sets of 11 two-alternative multiple-choice items) were the dependent variables. The same measures on a common instrument (36 five-alternative items), given to all *Ss*, were used as the respective covariates for the criterion variables. The Sheffé technique for determining confidence intervals was used to investigate the contrasts of interest. 60 *Ss* were randomly selected within each grade level. After administering the common instrument, the six experimental instruments were randomly assigned to *Ss* and administered.

The analysis of covariance for the time criterion resulted in no significant stem, FCDS, first or second order interaction effects. However, the results for score indicated significant stem, FCDS, and stem by FCDS interaction effects. All tests were at the .05 level. The results of the Sheffé analyses indicated that the positive items were significantly easier than the negative items, and within each stem mode, increasing the FCDS significantly decreased the mean score. An additional finding showed that the reliabilities of the negative instruments were similar to the reliabilities of the positive instruments, within comparable grade levels.

The results imply that if a subset of negative items is substituted for the corresponding subset of positive items, placing the negative subset at the end of the test would make the test less difficult than imbedding the negative subset among the positive subsets. The substitution of a subset of negative items for a subset of positive items may be made without fear of vitiating the reliability of the original set of items.

A COMPUTER APPROACH TO THE GRADING OF HISTORY EXAMINATION, *John F. McManus*, Southwest Regional Laboratory for Educational Research & Development; *Ellis B. Page* and *Dieter Paulus*, University of Connecticut

A design model was developed to determine how well a computer, using actuarial strategies, could evaluate college history answers written in response to short-essay identification items.

The data for the study comprised responses written by college freshmen and sophomores on a final examination in the History of Western Civilization. The answers averaged about 50 words in length, and there were approximately 100 responses for each of four items. The criterion for the study was the average grade assigned to each of the responses by a group of eight graduate students in history.

The predictor variables were of three types: (a) isolate words that discriminated between the good and the poor responses; (b) an interactive variable determined by multiplying the length of the answer times the total number of "key" words in that answer; and (c) two variables that had proven to be good predictors of style grade in Project Essay Grade, conducted at the University of Connecticut.

Eight analyses were performed, two for each response, once using the Type A and Type B predictors, and once using the A, B, and C predictors. Seven of the eight analyses revealed an empirically cross-validated multiple correlation coefficient of higher than .60 between predictors and human grade. A separate study also showed that the computer was able to evaluate well those answers that were long, but irrelevant.

A comparison showed that the computer graded essays a little better than the typical, qualified human, when the average human grade was used as criterion. Improvement in variable selection techniques should make the computer grades as reliable as those of two or three humans taken together.

THE USE OF MULTIPLE SET FACTOR ANALYSIS TO STUDY ESSAY RATINGS, *James E. Carlson*, University of Pittsburgh

Horst (1965) has proposed four different procedures of multiple set factor analysis which might best be described as extensions of canonical analysis to more than two sets of variables. The task of these procedures is to derive a set of weights for each of m sets of variables, such that the m resulting linear composites will be related in some desirable way. This paper deals only with one of these procedures, the maximum correlation method, which has as a criterion for the derivation of the weights the maximization of the sum of the $m(m-1)/2$ correlation coefficients between the m linear composites. The method was applied to the scores assigned by 16 essay readers ($m=16$) who each read 103 essay papers and graded them on 6 variables: spelling, punctuation, word usage, grammar, sentence errors, and form.

It was found that the weighting system resulted in an average correlation coefficient (inter-reader reliability coefficient) of .82 when the average coefficient for the unweighted composite score was .76, but that this increased coefficient resulted from the assigning of relatively large weights to one variable and small weights to the others. Implications with respect to selection and assignment of essay readers were discussed, as well as other possible uses for the procedure in educational research.

Session 5.5 **Division C**

INDIVIDUAL DIFFERENCES NO. 2

David Merrill, Brigham Young University, Chairman

A CYBERNETIC MODIFICATION SCHEME FOR AN INSTRUCTIONAL SYSTEM, *Michal C. Clark*, Stanford University; *M. David Merrill*, Brigham Young University

Adapting instruction to meet the needs of the individual learner is currently a goal of many experimental educational programs. This paper presents a modification scheme to allow a cybernetic instructional system to adapt itself to each individual.

This modification scheme assumes an aptitude by instruction interaction. It is assumed that in a group of students, different individuals will react differently to a given instructional program. Conversely, it is assumed that performance relevant to a set of objectives can be optimized by prescribing different instructional procedures to individuals differing in certain learning characteristics.

In this modification scheme a given type of objective is assumed to have a unique set of presentation and evaluation procedures. It is further assumed that these procedures are appropriate for all individuals but that the various parameters of these procedures have values which differ for different individuals. The proposed scheme describes a statistical sampling process for selecting the parameter values to be used for a given individual in a specific presentation. These modification procedures further specify mathematical operators which change on the basis of student response and, over several presentations, eventually select and stabilize on those parameter values which are optimal for the individual being instructed. In a very real sense this procedure allows the instructional system to learn to teach the individual. The parameter most frequently varied by programed instruction and CAI is pace. This scheme varies not only pace but a host of other variables dealing with the content of the presentation.

The paper to be presented will briefly describe the theoretical principles underlying the modification scheme, will list and illustrate at least one set of parameters, and will describe possible ways to implement the system using both CAI and a manual scheme for teacher use in the classroom.

AN EXPLORATORY INVESTIGATION OF THE CARROLL LEARNING MODEL AND THE BLOOM STRATEGY FOR MASTERY LEARNING, *Margaret Wang* and *C. M. Lindvall*, University of Pittsburgh

This paper describes results of an empirical investigation of relationships among variables of the Carroll learning model as they operate in individualized instruction. Analyses were made for six different samples of students studying units in elementary school mathematics under Individually Prescribed Instruction (IPI).

The Carroll model suggests that if each pupil is allowed the time he needs and is given instruction fitted to his requirements, every pupil can achieve mastery. The model further states that if the quality of instruction, the pupil's ability to understand instruction, and pupil perseverance are optimized for each pupil, then the time needed to achieve mastery will be a function of pupil aptitude.

In the current study efforts were made to adjust instruction to the needs of each student: each student was given the study time he required, and every student achieved tested mastery of the unit. One or more measures for each of the Carroll variables was obtained for each student. Included were four different rate measures and two measures of aptitude (non-verbal I.Q. and rate of learning in mathematics during the previous year). Data were analyzed using simple, muitiple, and canonical correlation.

Although relationships were not large or consistent over all units, the data do suggest that aptitude (as measured by previous rate) is a major factor in determining current rate of learning. This was indicated by the multiple and canonical regression analyses involving the weights for predicting rate from measures of the other Carroll variables. Where significant multiple correlations were obtained, structure-R's associated with aptitude ranged from .477 to .743. The analysis of results also indicates the usefulness of the Bloom strategy for evaluating the important components of a model for individualized instruction.

A STUDY OF THE EYE-MOVEMENT RESPONSES OF CHILDREN FROM VARYING GRADE AND INTELLIGENCE LEVELS TO DYNAMIC STIMULUS MATERIALS, *Willavene Wolf* and *Manfred Knemeyer*, The Ohio State University

The purposes of the study were to identify types of eye movements and to examine differences in eye movements responses of subjects of different grade and intelligence levels to various types of television materials. Twenty-seven students from grades 7, 9, and 11 served as subjects for the study. These students represented several intelligence levels. The stimulus materials presented to the subjects included video tapes of (1) an educational presentation with the narrator absent, (2) an educational presentation with the narrator present, and (3) a commercial.

Data were in the form of film records with the stimulus materials overlaid with the eye markers of subjects. These data were readout by computerized film reading equipment providing information on the position of the subject's eye markers from frame to frame.

Analyses of the data included (1) identification of the various types of eye movements made by the subjects; (2) the use of these types of eye movements as the dependent variables for the determination of the differences between subjects of varying age and I.Q. levels on the three stimulus presentations; (3) the examination of differences among subjects from varying age and I.Q. levels on their clustering patterns.

These analyses indicated that significant differences existed on the various types of stimulus materials, that one eye movement index was related to age and intelligence, and that differences did exist on the clustering patterns of subjects.

METHODOLOGY FOR EYE-MOVEMENT STUDIES IN DYNAMIC BI-DIMENSIONAL FIELDS, *Manfred Knemeyer* and *Willavene Wolf*, The Ohio State University

The purpose of this paper is twofold: (1) to compare experimental systems for recording eye movements of subjects while they view stimulus materials and to describe in detail the one used at The Ohio State University for recording eye movements to dynamic bi-dimensional fields; (2) to describe the system used to readout eye-movement data from films.

The experimental system to be described was designed for eye movement research in television viewing. The components of the system are (1) a corneal reflection system which is helmet mounted with remote controls for calibration purposes; (2) a vidicon television tube, also helmet-mounted; (3) vidicon controls; (4) a special monitoring system;

and (5) a recording camera. These components are integrated into a system which permits observation, monitoring and correction from a control room immediately adjacent to the experimental room. The basic component of the system is a modified head-mounted eye camera.

Data recorded using this system consisted of film frames with the picture that the subject is seeing overlaid with the subject's eye marker. The equipment for readout of these data consisted of a computerized film reader. Problems in developing a program for readout of the eye movement data, as well as problems encountered during the readout of the data, will be detailed.

Session 5.7 **Division C**

LEARNING FROM DISCOURSE: MATHEMAGENICS NO. 1

Dennis Anderson, Michigan State University, Chairman

THE EFFECTS OF QUESTIONS ON SHORT AND LONG TERM RETENTION, *Gerald L. Natkin* and *Elizabeth Stahler*, Bucknell University

The purpose of the experiment was to determine the effect of introducing relevant questions in prose materials, on long- and short-term retention. Recent research by Kleinsmith and others suggests that if questions cause arousal (GSR) increments, they should lead to a reminiscence effect. This effect should vanish if Ss are overexposed to questions, since the arousal response diminishes with stimulus repetition.

Ss for the study were fourteen male and female undergraduate students at Bucknell University. They were randomly assigned to either group (A) pre-exposure, or (B) no pre-exposure.

Students read two sequential 2500 word passages adapted from standard biology textbooks. The first dealt with the structure of plants, while the second covered ciliates. Group A received one question after each 100-word segment of the plants booklet; group B read the plants booklet without questions. Both groups read the plants booklet for fifteen minutes. Thus, group B should have been more aroused by questions at later points in the program than group A. All Ss then received and completed the ciliates passages, which had a question after each page. They then took an immediate post-test on ciliates. One week later all Ss took a delayed post-test on ciliates.

The results were as predicted. The mean for group A decreased over time, while that for group B increased (F=22.088, df=1, 12).

These results strongly support the hypothesis of a reminiscence effect. It may prove possible to apply similar methods to improve retention from prose reading, providing that methods for maintaining stable arousal responses can be developed.

EFFECTS OF ADJUNCT QUESTIONS, PRETESTING, AND DEGREE OF STUDENT SUPERVISION ON LEARNING FROM AN INSTRUCTIONAL TEXT, *H. W. Gustafson* and *David L. Toole*, Bell Telephone Laboratories, Incorporated

Fifty-eight high-school seniors were paid to engage in a self-paced study of a 23,000-word expository text on computers. The text, though "unprogrammed," was fully self-instructional, consuming a mean study time of 4.9 hours spread over several days.

A 2x2x2 design was used involving the following variables: (1) 53 adjunct questions *vs.* no adjunct questions, (2) a 24-item pretest *vs.* no pretest, and (3) scheduled, supervised study at school *vs.* independent, unsupervised study mostly at home. Each student reported for posttesting whenever he was ready and was posttested 24 hours later. A retention test distinct from the posttest was administered 7 to 13 days afterwards.

The posttest consisted of the 24 pretest items, plus the 53 adjunct questions, plus an additional 44 new items.

No effects were detected on any measure (including study time) for either variable (1) or variable (2). In other words, neither pretesting nor supervision of study appeared to have any influence on achievement.

Students exposed to adjunct questions, on the other hand, achieved significantly better than those not exposed, but only with respect to that subtest of the posttest comprised of the adjunct questions themselves. It thus appears that the adjunct questions, though generating strong question-specific effects, failed to produce the more general effects on learning observed in recent experiments by others. We surmise that the general beneficial effects elicited in investigations elsewhere are attributable to the experimental requirement that the subjects read the designated instructional material only once, which would seem to make the findings of such investigations more germane to sequential reading than to the process of painstaking study. In the study process, where rereading and review are prime ingredients, it may be that adjunct questions lose their potency except as emphasizers of the specific information with which they deal.

THE EFFECTS ON LEARNING OF THE POSITION OF AN ORGANIZER, *Daniel J. Bauman* and *Gene V. Glass*, Laboratory of Educational Research, University of Colorado

The purpose of the study was to determine the effect which position of an "organizer" (before or after a lesson) has on achievement in learning a lesson. The hypothesis was tested that S's receiving an organizer after a lesson would perform better on a test over the lesson than either those receiving an organizer before a lesson or those receiving no organizer.

Two sets of material were prepared, one on product-moment correlation and the other on hypothesis testing. Two classes, each consisting of 21 students, in introductory statistics provided the S's—one class for each experiment. Each class was randomly divided into three groups of seven S's each. A control group received the lesson only; a second group (pre-organizer) received the organizer followed by the lesson; the third group (post-organizer) received the lesson followed by the organizer. All members of each class then took the same twenty-item test. The organizer consisted of abstract, general material intended to function as an "organizer" of the material to be learned in the manner indicated by Ausubel (1960).

The inferential statistical analysis of the data was by means of two planned, orthogonal contrasts.

In both experiments, the post-organizer group scored significantly higher than the pre-organizer group. Also in both experiments, the average level of the two "organizer" groups was not significantly higher than that of the control group. The results are interpreted in terms of Rothkopf's mathemagenic behavior.

MOTIVATION AND MATHEMAGENIC BEHAVIORS: TOWARD A THEORY OF INSERTING QUESTIONS IN INSTRUCTIONAL MATERIAL, *John Romney Pyper*, Purdue University

This paper presents a theoretical analysis of research concerned with the applied instructional task of improving the learning that occurs from the study of didactic textual information. Task analysis suggests that a principal component of such a task is the manipulation of the nature and intensity of the Ss' attentional processes (mathemagenic behaviors). The procedure for manipulating attention, to be reviewed, is that of inserting questions in textual materials. Such an approach has been studied in relation to the learning occurring from films (Michael and Maccoby, 1961 and Maccoby, Michael and Levine, 1961). More recently the effects of inserting questions in written textual material have been rather intensively investigated by Hershberger (1964), Hershberger and Terry (1965), Rothkopf (1965 and 1966), Rothkopf and Bisbicos (1967), Frase (1967 and 1968), and Pyper (1968).

The research results generally indicate that questions inserted after (rather than before) sections of relevant text facilitate not only the learning of the information directly tested by the questions, which would be expected, but also the learning of the information untested during the study of the text. Further research indicates that the effect of inserted questions is a function of difficulty, novelty, and content as well as placement, of the questions. These characteristics interact with task set and novelty of the task in determining the degree of knowledge of task (KOT) and task accomplishment incentive (TAI). KOT and TAI determine Ss' perceptual and rehearsal activities (mathemagenic behaviors) in relation to the text. The nature of the mathemagenic behaviors determines the extent of learning. This cognitive task analysis approach is contrasted with an S-R analysis of a restricted aspect of the task.

Session 5.10 **Division D and NCME**

EMPIRICAL TEST RESEARCH

J. Thomas Hastings, University of Illinois, Chairman

THE PRINCIPAL COMPONENT STRUCTURE OF A PRODUCTIVE THINKING BATTERY, *James L. Bavry* and *Herbert J. Klausmeier*, University of Wisconsin

The purpose of this study was to investigate the interrelationships among some measures of hypothesized productive thinking abilities. Most measures analyzed were produced by members of the Productive Thinking Program at Berkeley, California, according to a measurement schemata given by Covington (1968). Four tests (providing 12 measures) from the Torrance Tests of Creativity and measures of intelligence and achievement were included in the analysis. It was undertaken (1) to see if the dimensionality of the measures was consistent with Covington's measurement schemata, and (2) to determine the relationship of the measures to each other.

Data for 119 control subjects (59 males, 60 females) were collected as part of an experimental study evaluating the Productive Thinking Program materials (Olten, et. al., 1967). Fifty-four of over 150 measures collected were chosen for analysis. The correlation matrix for these 54 measures was factored according to two procedures: (1) the incomplete principal component method, and (2) the Harris R-S^2 solution. In both cases varimax rotations were performed. Method one produced 17 factors; method two produced 34 factors, two of which were null.

General conclusions were that (1) the measures in the battery cover an extensive area independent of intelligence and achievement, (2) different measures obtained from the same test loaded together, and (3) several measures important in Covington's scheme followed a pattern consistent with his theorizing. Dimensions corresponding to general abilities either to produce ideas in quantity or to produce ideas rated as high quality were not observed.

The results of the present study suggest that more attention should be paid to the specification of components of variance attributable uniquely to the form of productive thinking measures. Validational studies would provide for both a more rigorous testing of the hypotheses of Covington and a specification of variance components attributable to instruments and traits.

MINIMIZING PROXIMITY ERRORS IN THE SEMANTIC DIFFERENTIAL, *Robert B. Kane*, Purdue University

The purpose of this research was to ascertain effects of manipulating sources of proximity error on response patterns among *S*s completing semantic differential (SD) questionnaires.

A computer program was created to generate SD's such that, for a given set of questionnaires, any or all sources of proximity error may be minimized. The sources include concept order, scale order, and adjective order within scales. Cyclic Latin squares selected by a heuristic strategy served as generating keys for concept and scale orders that minimize proximity error. SD's containing as many as 22 concepts, each with as many as 22 scales, may be produced. A pseudo-random digit generator served as a key for determining adjective order within scales. The program may be modified to produce other types of questionnaires or tests with proximity errors minimized.

Three experiments were conducted. In Experiment I, responses to questionnaires having all three sources of proximity error minimized were compared with responses to questionnaires having all three orderings invariant. In Experiment II, questionnaires with all three orderings invariant were replaced by questionnaires having four concept orders, but with scale and adjective order invariant. In Experiment III, the questionnaire formats were (1) only adjective order within scales varied and (2) four different concept orders with scale and adjective orders invariant. One hundred fifty *S*s were selected randomly and assigned to treatment groups randomly. Each SD contained nine concepts, each with 14 scales. For each concept within each treatment, a 14x14 matrix of intercorrelations was factored and rotated.

No differences in rotated factor structures, SD factor scores, or response consistency attributable to item orders were observed in any experiment. Subject to generalizability constraints, it appears that proximity error does not induce significant differences in *Ss'* response patterns with SD's. Thus *Es* need not be concerned about proximity errors when administering the SD.

MEASURING ATTITUDE BY AN EXTENSION OF THE SEMANTIC DIFFERENTIAL, *P. R. Merrifield*, New York University; *M. Arnoff*, Kent State University

Osgood and others have reported the semantic differential technique for the pursuit of meaning; differences among the concepts used as referents have been noted. The present paper reports a method that capitalizes on such differences to establish measures of attitudes of individuals to underlying concepts.

The first of three phases of the method is familiar: consider all responses for all concepts as replications of the set of adjectives; factor the intercorrelations among the adjectives; interpret the rotated orthogonal factors in such terms as activity, potency, warmth, and acceptability.

Phase II requires the computation of factor scores for each individual's response to each concept. These data are grouped by concept, and the mean and variance of each factor score obtained for each concept. Three kinds of comparisons are made: location of concepts in the common-factor space, intra-concept profile in terms of all factors, and inter-concept relations for each factor separately. These comparisons involve multidimensional distance measures and the t- and F-distributions.

In Phase III, relations among the concepts with respect to a single factor are analyzed. Factor scores from Phase II are rearranged into a concept-by-individual matrix, and the inter-concept correlations computed and factored. The correlations between concepts reflect the degree to which those individuals who judge Concept A to be more "active," for example, also judge Concept B to be more active, relative to the means of active-ness for the two concepts. The common ground of each group of concepts that cohere as a factor is interpreted as an underlying concept toward which each individual's more basic attitude can be measured.

Results of the application of the method to responses of middle-school age children to adjectives and concepts from the domain of societally relevant issues will be reported.

MEASUREMENT OF GRADUATE SCHOOL ENVIRONMENTS, *James M. Richards, Jr.*, American College Testing Program; *Richard Seligman*, University of California, Los Angeles

For several years, researchers have devoted considerable effort to discovering the ways in which colleges influence their students. As part of this effort, a number of techniques for measuring college environments have been tried. A number of these have a common difficulty, namely that they are based on characteristics of students such as their choice of a major field or their responses to a questionnaire. Because they confound student characteristics with college environment characteristics, such techniques produce many problems for subsequent investigations of college influence on students.

A recent study suggested that a promising technique for measuring undergraduate environments is to group the faculty and the curriculum into a profile according to a psychological classification scheme for major fields. Such profile scores measure the college environment independent of student characteristics, appear to reflect differences in the emphasis of different kinds of colleges, are fairly reliable, and are related in meaningful ways to other measures of the college environment. The present study extends this technique to the measurement of graduate school environments.

The sources of data were 1968 graduate school catalogs from 86 institutions. The basic procedure was to count the number of graduate courses and faculty that fell into each of six types of major field—Realistic, Intellectual, Social, Conventional, Enterprising, and Artistic—based on Holland's theory of vocational choice. Transformation to normalized standard scores yielded profiles for each graduate school. Like all profiles, it can most appropriately be analyzed in terms of three components: elevation, scatter, and shape.

The profile scores were correlated with CUES, environmental scores published by Astin, and Cartter's estimates of quality of Ph.D. training. Correlations with other environmental measures generally support the construct validity of the profile scores. Quality is most related to the size of the graduate school. Overall results suggest this is a promising approach worthy of further study.

Session 5.11 **Division E**

COUNSELOR CHARACTERISTICS AND COUNSELING INNOVATIONS

Lester M. Snyder, Arizona State University, Chairman

EXPLORATION WITH A COMPUTER-ASSISTED OCCUPATIONAL GUIDANCE PROGRAM, *Joseph T. Impellitteri*, Pennsylvania State University

Based upon the needs of today's youth to explore the world of work and the increasing importance of vocational guidance in their development, the Computer-Assisted Career Exploration System was developed. The purpose of the system is threefold: to provide an easily updated individualized occupational information retrieval system; to develop through an essentially heuristic approach a process whereby youth could develop their own individualized frameworks of the occupational structure; and to provide an experience for youth to acquire, by simulated practice, operational strategies in relating their abilities and interests to occupational opportunities.

The system was tested in two separate trials. The first trial, approximately ten weeks in length, was held in the spring of 1967. The youngsters scheduled for the system experience were 72 ninth grade boys who had indicated that they were interested in selecting a vocational or technical course of study the following year at the senior high school. The second field trial involved the scheduling of 140 ninth grade boys between October, 1967 and May, 1968. These boys also were selected on the basis of their having indicated a preference for a vocational or technical course of study the following year.

Based upon the results of the field trials, it may be concluded that the computer-assisted career exploration system is an effective and feasible device in assisting youth to explore occupational opportunities. It also may be concluded that the lack of a sound theoretical foundation in vocational psychology impeded the investigators in drawing any implications of the study with regard to such factors as vocational maturity. It is thereby recommended that future efforts in this area be focused upon a theoretical as well as a problem orientation.

CLIENT BEHAVIOR CHANGE AND THE PSYCHOLOGICAL CHARACTERISTICS OF THE COUNSELOR, *Jean C. Waterland* and *Ernest Spaights*, University of Wisconsin-Milwaukee.

The evidence suggests that counselor candidates who operate as effectively as possible with their clients seem to possess personal characteristics which other counselor candidates appear to lack. The purpose of this study was to determine what differences existed between effective and ineffective counselor candidates as determined by their ability to effect client behavior change.

Fifty counselor candidates participated in the study over a period of three semesters. The program was designed to assist counselor candidates develop their verbal and non-verbal counseling skills by using a multiple approach. Each semester the counselor candidates completed a battery of psychological instruments, participated in a programmed approach to sensitivity training, counseled clients throughout the semester, had scheduled supervisory sessions, and participated in videotaped counseling sessions. Two supervisors independently rated each counselor candidate on a continuum of effectiveness by using pre-established criteria which focused on client behavior change. There was seventy-two per cent agreement between raters.

The performance of the effective and ineffective counselor candidates on the psychological instruments was compared. Although no significant differences existed between the two groups, the effective counselor candidates demonstrated a tendency to be more achievement oriented, had less need for sympathy, felt less guilty, enjoyed more social activities, were more accepting and non-judgmental, were aware of their anxiety, and were less interested in economic values than the ineffective candidates.

DETERMINING THE EFFECTS OF OVERT AND COVERT RESPONDING TO SOCIAL MODEL COUNSELING PROCEDURES: A STUDY WITH SHY ELEMENTARY SCHOOL CHILDREN, *David T. Antonioni*, University of Wisconsin-Madison.

The general purpose of this study was to develop and experimentally test counseling techniques that can be used by counselors to assist shy elementary school children to participate more in classroom discussions.

This study investigated the relative effects of using a film social model accompanied by one of two types of rehearsal (overt and covert responding) after exposure to the film model. Overt rehearsal consisted of actively recalling the film-model content plus role-playing the model's responses. Covert rehearsal consisted of "thinking" about the model's responses only. The two dependent variables which the experimental procedures sought to promote were: (a) retention, i.e., amount of film model content the observer can recall, and (b) performance, i.e., amount of voluntary classroom participation responses. Male and female fourth, fifth, and sixth grade students were utilized for the research project.

Three hypotheses were formulated for testing: (1) Subjects who observed the film model and who engaged in either overt or covert rehearsal would exhibit significantly greater frequency of voluntary verbal classroom responses than would similar subjects who received no counseling; (2) Subjects exposed to overt rehearsal would emit a greater frequency of voluntary verbal classroom responses than would subjects exposed to covert rehearsal; and (3) Subjects exposed to overt rehearsal would recall significantly greater content of the film model interview than would subjects exposed to covert rehearsal.

Hypothesis one was supported. Subjects exposed to the social model treatments demonstrated significantly more voluntary participation than did similar subjects receiving no counseling. There was no significant difference between overt and covert rehearsal groups.

That learning does take place by exposure to models was evident. The present study has demonstrated that a film model can bring about new behavior with shy students. Additional research is needed, however, to determine what types of shy students identify with what types of models and perform better when supplementary social model counseling techniques are used.

Session 6.3 **Division C**

VERBAL LEARNING NO. 1

Robert E. Davidson, University of Wisconsin, Chairman

THE EFFECTS OF A MEDIATIONAL DEVICE ON PAIRED-ASSOCIATE LEARNING: LEARNING AND RETENTION OF SIXTH GRADE STUDENTS, *William E. Loadman*, Michigan State University

The purpose of this investigation was to determine the influence of different levels of strategies for paired-associate (P.A.) learning on accuracy, speed of learning, and long term retention.

The experiment was conducted on three randomly selected groups of sixth grade students. Each group of approximately 25 middleclass children were given a practice P.A. task; this was immediately followed by the experimental task where the groups were randomly placed in a control or experimental section. The control Ss were instructed to learn the task the best way they could. The E_1 group was given a high level strategy on trials 1 and 2 (the P.A.'s were placed in sentences and read aloud to the group) and instructed to learn the list. The E_2 group was given examples of high level strategies, and during trials 1 and 2 they were instructed to form their own high level strategies. The list consisted of 8 dyads of low meaningful stimulus and high meaningful response words. A recall trial was presented after each of the 8 learning trials. One week later a measure of retention of the P.A. list was obtained.

There were no significant differences between groups on the practice task. On the experimental task, the E_1 group had significantly more words correct and learned the list significantly faster than the control or E_2 group. On the retention task, the E_1 and control groups were significantly different from the E_2 group but not from each other. It appears that speed and accuracy of learning are facilitated by invoking a high level strategy on the group, but not by having the group form their own high level strategy. Forming one's own high level strategy also appears to impede retention. This suggests that foreign language vocabulary learning may be best accomplished through teacher initiation of high level strategies.

VERBAL VS. PICTORIAL FACILITATION OF PAIRED-ASSOCIATE LEARNING, *Janice Freeman Adams* and *Robert E. Davidson,* Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin

Verbal and pictorial mediational links have been shown to facilitate the learning of common-noun paired-associates (PAs). The facilitating effect of verbal strings seems to vary with the part of speech of the connective in the verbal string, e.g., conjunctions, prepositions and verbs produce different rates of learning. The facilitating effect of pictorial linking seems to vary with the spatial relationship between the paired objects, i.e., whether the objects are placed side by side or in some other relationship to one another (e.g., under, in, on).

This experiment was designed to compare the relative facilitating effects of prepositional and conjunctive strings when they were accompanied by pictures of the pair that were side by side or pictures of the pair that were translations of the prepositional strings.

Second grade children served as subjects in an independent group 2 X 2 factorial design. Over two study-test trials the subjects learned 20 picture pairs in a recognition PA task. The response measure was the total number of correct recognitions over both trials.

The results indicated that prepositional strings, or pictorial linking that is a translation of prepositional strings, or both greatly facilitates the learning of paired-associates. There was evidence that a prepositional link was more effective than a pictorial translation of a prepositional link. This may be interpreted as support for the idea that the covert processes that underlie both verbal and pictorial facilitation are verbal in nature.

THE FREE RECALL OF WORDS FOLLOWING SEMANTIC CONDITIONING, *Betty C. Porter* and *Robert E. Davidson,* Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin

Associative interrelationships among words are well documented in the literature; these interrelationships predict that related words will occur in common verbal environments. The purpose of the present study was to investigate free recall of words presented within a semantic conditioning paradigm.

Via tape recorder, 30 *S*s heard 40 words throughout three conterminous phases of the conditioning paradigm (habituation, conditioning, generalization). The word stimuli were drawn from seven associative structures which were derived originally from a factor analysis of the associations to 120 words. The word "violin," from a primary factor cluster, served as the *CS*. White noise was the *UCS*. Upon completion of the three phase experiment, *S*s were asked for their free recall of the words.

Word recall was a function of cluster interrelationships and frequency of repetition. For example, words from the primary (signal) cluster were recalled more often than words from more remote generalization clusters. Both frequency and patterning of word recall suggest that it is possible to describe objectively the semantic "fields" that surround words.

COGNITIVE FACTORS IN SEMANTIC CONDITIONING, *Richard T. Walls*, West Virginia University; *Francis J. Di Vesta*, The Pennsylvania State University

This experiment was designed to determine the effects of frequency of reinforcement, repeated evaluation of stimuli, and cognitive correlates upon classical semantic conditioning of preference. Previous experimental results, in which the effects of conditioning (reinforcements) were inseparable from the possible effects of repeated ratings, found a linear increase in preference for a critical stimulus over four rating days alternated with three conditioning days.

The present design orthogonally crossed two levels of stimulus conditions (same stimuli throughout and different stimuli on each new conditioning day) with four levels of rating schedules (days 1, 3, 5, and 7; days 3, 5, and 7; days 5 and 7; day 7). That is, *S*s at the fourth level rated only on the final day and performed a filler task on days 1, 3, and 5. A conditioning session consisted of a first grade *S* spinning a spin-wheel apparatus, on which were printed Greek letters, until 20 reinforcements (marbles) had been received for a preselected critical stimulus. The rating procedure, from which a preference score was derived, consisted of *S*s assigning each of seven positive evaluative and seven negative evaluative adjectives to any of six stimuli among which was the critical stimulus.

The comparable condition reproduced, exactly, the results of the previous work; however, as hypothesized, in addition to the strengthening effect of reinforcements, preference was a direct function of the number of ratings. It was proposed that this effect was partially due to increased perception of elements of the demand characteristics of the experiment. Assessment of *S*s' reinforcement and behavioral hypotheses incorporated a role reversal procedure in which *S* attempted to appropriately administer reinforcements as E played the spin-wheel game. The role exchange procedure was viewed as a potentially fruitful means for assessing awareness through direct behavioral measures.

Session 6.4 **Division C**

PRESCHOOL TRAINING PROGRAMS NO. 1

Robert Boger, Michigan State University, Chairman

AN INVESTIGATION OF MONTESSORI VS. CONVENTIONAL PREKINDERGARTEN TRAINING WITH INNER CITY CHILDREN: AN ASSESSMENT OF LEARNING OUTCOMES, *Barbara Berger*, Center for Urban Education, New York City

The research design provided for replication by using two independent comparison studies of children undergoing Montessori training vs. conventionally taught children. Each research setting included experimental (Montessori) and control (conventional) classes, involving three Montessori and two control teachers. The sample consisted of 92 Puerto Rican and Negro children, chiefly four year olds, but also some three year olds.

This investigation of training effects was based on the first year of schooling. It included an assessment of perceptual and conceptual functioning, as well as an assessment of cognitive styles specifying self-regulatory, autonomous problem-solving orientations. Two different test batteries were employed, each individually administered at the end of the school year.

Results in each instance showed some significant differential effects of training. These were most pronounced with respect to cognitive style differences, where the Montessori model was consistently more effective than conventional approaches in reinforcing autonomous problem-solving orientations. Thus, the data indicated Montessori's major strength to be the psychological impact of the teacher-child transactions in a Montessori environment, rather than the achievement impact. However, the Montessori approach was also uniformly more efficient in upgrading perceptual discrimination skills for the poorest achievers in this assessment area, at both age levels investigated, a finding which substantiates the merits of its highly programmed perceptual training techniques.

No comparable systematic trends, however, were apparent with respect to conceptual functioning, where the data did not favor one teaching approach or the other. Here the training picture indicated more variable outcomes for Montessori teaching as a function of individual teacher styles and instructional emphasis.

AN EXPERIMENTAL PROGRAM DESIGNED TO INCREASE AUDITORY DISCRIMINATION WITH HEAD START CHILDREN, *Charlotte Ann Brickner*, University of Colorado

The purpose of this study was to test two planned training sequences designed to increase auditory discrimination. Two treatment groups and a control group were randomly selected, following blocking on auditory, visual, and auditory-visual learning style, from a population of 98 Head Start children. Treatment I consisted of eighteen programmed training sequences, Treatment II consisted of eighteen narratives of an appropriate level which were read to the children, and the control group received no treatment. Both treatments were presented on magnetic tapes.

Analysis of data was done utilizing analysis of variance and analysis of covariance and a three dimensional factorial design.

The criterion measure was designed to test three aspects of auditory discrimination: sound discrimination, verbal discrimination, and following directions. Both treatments showed consistent significant ($p < .01$) differences over the control group following treatment on all aspects of the criterion measure. Treatment I was significantly superior ($p < .01$) to Treatment II for verbal discrimination and following directions.

If auditory discrimination is a factor which may contribute to a child's ability to learn to read as has been suggested by some authors, then increasing this ability to discriminate may have some effect on early education. The results of this study indicate that a child from this population can be taught to attend to individual sound stimuli. Attending to individual stimuli is prerequisite to categorizing and classifying classes of sound stimuli, the definition, for the purposes of this study, of auditory discrimination.

MOTHERS AS TEACHERS OF THEIR OWN PRESCHOOL CHILDREN: THE INFLUENCE OF SOCIOECONOMIC STATUS AND TASK STRUCTURE ON TEACHING SPECIFICITY, *Jere Edward Brophy*,

The University of Texas at Austin

This research investigated informational specificity as a variable in the teaching behavior of mothers interacting with their preschool children. Variation in specificity is studied in relation to differences in socioeconomic status (SES) and in certain structural characteristics of different components of the teaching task.

The data were assembled by applying a specially developed coding system to typescripts from an observed and tape recorded mother-child interaction on the Block Sorting Task (Hess and Shipman, 1965). A total of 137 urban Negro mothers and their four-year-old children, divided into four SES groups, were included.

As hypothesized, specificity was found to be associated with SES. Higher status mothers tended to give more complete information (explicitly verbalize relevant concepts or labels) than lower status mothers at parallel points in the task, and they also made greater use of non-verbal "focusing" techniques which increased the salience of relevant stimulus attributes and helped the child to associate the verbal labels with their specific physical referents. Specificity also varied with task structure, being higher in post-response feedback than in pre-response instructions or orientation and higher following errors than correct responses.

In combination, the data suggest that SES differences in maternal teaching behavior are not describable in terms of two contrasting, crystallized "styles," but instead represent variation from limited, primarily reactive teaching to more diversified, proactive teaching. Specificity is highest (and SES differences are minimized) when task structure and preceding events combine to elicit specific information from the mother as a reactive response. In the absence of such external press, specificity will occur more as a proactive maternal teaching strategy. Here specificity occurs least consistently and variation related to SES is maximized.

HETEROGENEOUS VS. HOMOGENEOUS SOCIAL CLASS GROUPING IN PRESCHOOL PROGRAMS, *Sarah D. Hervey* and *Robert P. Boger*, Michigan State University

Programs for disadvantaged children tend to single them out and isolate them from the rest of society as the special program is conducted. Such isolation is often necessary if the program is to have maximal impact upon the group it serves by focusing all its resources on that group. But such isolation must also be balanced against the adverse effects that isolation might have upon the very individuals a program seeks to serve. There is ample evidence to support the contention that children learn from one another in any social setting, and it is reasonable to hypothesize that disadvantaged children would learn from the advantaged children with whom they interact in school. The present study tested this hypothesis.

The variables studied were cognition, language development, and socialization. It was hypothesized with regard to the latter that disadvantaged children would model their behavior after that of their more advantaged peers.

Three Head Start classes were involved in the study. The experimental and control groups were matched ethnically and randomly assigned within ethnic groups. The Wechsler Primary Scale of Intelligence (WPSI), the Cincinnati Autonomy Test Battery (CATB), Kansas Social Interaction Observation Procedure (SIOP), videotape recording, and sociometric and self-concept measures were employed. An F test on the pre-post gain scores was used in analyzing the data.

The results indicated support for the major hypotheses, particularly in areas of social-emotional functioning.

Session 6.7 **Division F**

TRANSFORMATION OF THE SCHOOL REVISITED

Warren Button, State University of New York, Buffalo, Chairman

PROGRESSIVE EDUCATION: THE PRE AND POST WORLD WAR I EMPHASES, *Patricia Albjerg Graham*, Barnard College, Columbia University

The progressive education movement is now commonly described as beginning in 1876 and ending in 1955. Its intimate relationship to the progressive movement in American politics and to the New Deal is frequently asserted. A major historiographic discussion currently concerns the degree of continuity between American political progressivism and the New Deal. Similar questions are appropriate to raise regarding the reforms in education during the periods before and after World War I.

The argument for examining the pre and post World War I periods separately rests largely upon the different emphases of the reform movement in education during those times. Prior to 1918, a characteristic of the movement was its concern with problems of urbanization, industrialization, and immigration, all developments which most influenced the lower socioeconomic classes. Generally the school was asked to enlarge its functions to become a leader in community welfare.

After 1918, the characteristics of the movement changed. The educational reformers directed their principal attention in the twenties to releasing children's creativity, particularly in new, privately-supported schools; in the thirties to theorizing about the role of the school in the social order; in the forties to reorganizing the school program so that it was tied less to the traditional academic discipline and more to the interests and needs of children. The educational reformers of the post World War I period focused their interests upon schools and children largely in affluent circumstances.

Whether one decides to call these reforms all part of one movement, but one with considerably different emphases in its early and late stages, or rather two separate but related movements is perhaps not as important a question for a historian as a sociologist. Of greater interest to the historian, however, is recognition of the distinctive qualities of educational reform in the beginning of this century and again in the years between the world wars.

SOCIAL EDUCATION AND THE PROBLEM OF THE COMMUNITY IN AMERICAN LIFE, *Joel H. Spring*, University of Wisconsin

Major institutional changes in public education in the United States in the early part of the twentieth century can be explained in terms of a response to what was believed to be the breakdown of the community in American life. Among these changes were the development of vocational guidance, junior high schools, tracking, the use of psychology in the schools, student government, and an organized school social life. The purpose of my paper will be to place these educational developments, which I call social education, into the framework of what was a general reconsideration of the social organization of the United States.

This general ideological framework encompasses intellectual patterns that developed in the fields of philosophy, social psychology, and progressive politics between 1890 and 1920. The common strain of thought that runs through these fields was that the solution to the social problems faced by Americans could only be solved by finding a new basis for national unity. The general argument put forth was that early nineteenth century America maintained its cohesiveness through a community life based on face-to-face relationships and on the individual's awareness of the value of his labor to the total community. It was believed that the impersonality and demand for specialized economic roles that accompanied the urban-industrial society as it emerged after the civil war had destroyed community life and alienated the individual from the social system. The main thrust of the thinking of social psychologists like James Baldwin, Edward Ross, George Herbert Mead, and Charles Cooley, of philosophers like Josiah Royce, John Dewey, and Herbert Croly, and of Theodore Roosevelt's New Nationalism was that a new form of community life had to be found for the modern social order. It was hoped this new community life would allow for needed social specialization, but also that it would end individual alienation and create a new form of social cooperation.

I maintain that it was this pattern of thought that had the greatest impact on American public education in the early twentieth century. I consider this argument to be a revision and expansion of Cremin's *The Transformation of the Schools.* Cremin neglected the progressive ideology of Theodore Roosevelt's New Nationalism. It was this ideology that was reflected in the educational style of the American public schools and in the major developments in social thought in the United States.

CHILD-CENTERED PEDAGOGY IN THE 1920's, *Jonathan E. Burack,* University of Wisconsin

The treatment of what Cremin calls the "Sentimental" school of progressive educators which emerged just before and during the twenties has not, in my opinion, been adequate. This group has been passed off as a brief and rather irrelevant aberration from the healthy liberalism of the authentic progressive tradition. Naive romantics, they provide cannon fodder for critics of progressive education, and their excessive concern for a false individualism quickly passed with the more pressing and "real" concerns of the depression decades.

Whether or not one chooses to endorse the members of this group, a more thorough analysis of them will reveal that they represented something that was, and still is, quite important to American education. They were responding to very real failures in the philosophy and (more importantly) the practice of the more orthodox progressives of the period. More broadly, they were reacting to the growing power of mass technological society over the lives and very thought processes of individuals. The tendency of progressive thought to become nothing more than a benign form of manipulation and social control disturbed them and they groped (however unsuccessfully) for a new definition of real individual freedom. It is important to ask how they conceptualized this freedom, what problems they felt themselves pressed to solve, and why they failed (assuming, as I do, that they did fail).

A number of specific concerns are the focus of this study. These are: (1) the fact that so many of this school of progressives were women, (2) the apparent preference of the Freudians among them (Naumberg, for example) for Jungian psychoanalysis, (3) the centering of this group in New York, (4) their attitudes toward other groups of educators, and (5) their general social thought. I also hope to place their thought in the broader perspective of the process of institutionalizing of social services (psychoanalysis among the medical groups, social work, etc.). My thinking here has been influenced by more radical and contemporary interpretations of Freud (i.e., Brown and Marcuse) to the feeling that this initial "romantic" response to Freud may not have been without cause or justification.

RESPONSE, *James Wallace,* Reed College

Session 7.1 **Division B**

MODELS FOR CURRICULUM DEVELOPMENT

George A. Beauchamp, Northwestern University, Chairman

ONE MODEL FOR CURRICULUM DEVELOPMENT: PROBLEMS AND POSSIBILITIES, *Jack R. Fraenkel*, San Francisco State College

One of the chief concerns of the Taba Curriculum Development Project at San Francisco State College has been the development of more adequate curriculum theory. As part of our work, we have designed a curriculum model that

(1) includes an organization of, and relationships among, objectives, content, learning experiences, teaching strategies, and evaluative measures so that a system of teaching and learning is represented;

(2) appears generically applicable to a wide number of curricular areas (It has been used in biology, economics, drama, mathematics, and elementary social studies);

(3) has been used in curriculum development and teacher training in a number of different settings:

a) with different school systems—from rural to suburban to urban, from affluent to impoverished, and from those with ample to those with sparse professional resources;
b) in a wide variety of teacher training and informational situations—one to two week summer workshops, NDEA institutes of several weeks' duration, week-end sessions, year-long programs meeting on a once or twice monthly basis, hour-long briefing sessions at professional meetings, and "one-shot" consultancies with individual schools or districts;
c) in pre-service courses in curriculum and instruction;
d) in different curricular areas (e.g., biology, economics, drama, and mathematics).

These efforts have led us to a number of conclusions about the development and dissemination of curriculum innovations. These conclusions, together with the model, are presented; problems encountered in the development and implementation of the model are discussed; and alternative possibilities for dissemination are suggested.

EMPIRICALLY DEVELOPING CONCEPTUAL MODELS OF CONTENT AREAS, *David J. Pucel*, *Brandon S. Smith* and *Jerome Moss Jr.*, University of Minnesota

This was an exploratory study which investigated the efficacy of an empirical procedure to identify and generate the psychological structure of a content area possessed by a group of persons applying the content optimally. A further problem was to determine the sensitivity of the methodology.

The content area selected for the study was the technical knowledge necessary to perform as a radio-tv repairman.

A random sample of technical electronic words, representing technical concepts, was selected and presented to a small purposive sample of "flexible" and "inflexible" radio and tv repairmen in the form of a continuous response free-association instrument. Flexible workers were defined as being capable of performing a greater range of repair tasks than inflexible workers.

A special computer program was used to calculate the associative commonality among the response distributions of each group of workmen to all possible pairs of stimulus words. The resulting matrices of associative commonality measures for each of the two groups of workmen were then factor analyzed to reduce the matrices to simple structure.

The following results were obtained:

1. It was possible to generate a hierarchical associative structure for each of the two groups.
2. The associative measures of verbal behavior seemed sufficiently reliable (.78—.82).
3. The methodology was sensitive enough to differentiate groups of people performing the same task at different performance levels in terms of their psychological structure of the content area.

Although more research must be done to determine the validity of the methodology, the methodology appears to be sensitive and to be capable of empirically generating a hierarchical associative map of a content area. The map of the flexible workers (the optimal group) might be used as a model toward which curriculum developers may wish to build curriculums in an attempt to help students develop and learn the content interrelationships possessed by persons using the content optimally. The methodology also has direct implications for evaluating and diagnosing educational or conceptual weaknesses of individuals who are not performing tasks related to a content area optimally.

A MODEL FOR SYSTEMATIC RESEARCH AND DEVELOPMENT OF CURRICULUM PROGRAMS, *Wai-Ching Ho*, Educational Research Council of America

With the explosion of the new curriculums in the last decade, never before has the world of educational research raised so many questions and elicited so many new ideas as well as controversies concerning various aspects of curriculum development and evaluation. What is lacking is a continuous, systematic, and well-coordinated plan for the different stages in the research and development of a new program. It is the purpose of this paper to present a basic scheme to attack this problem.

The scheme encompasses 9 basic steps: (1) analyzing educational needs, surveying existing curricula, and soliciting experts' opinion; (2) formulating program rationales and objectives; (3) determining learning experiences and drafting tentative program units; (4) experimenting with the tentative units and collecting relevant information; (5) designing and developing the complete program; (6) implementing the program in the classrooms; (7) evaluating effectiveness of the program and conditions for its best usage; (8) drafting and experimenting with new or revised materials; and (9) repeating steps 5-8 for the revised and further revised editions of the program.

This model calls for a large number of well controlled, short-term experiments in step 3. It is contended that much valuable information (such as readability, logical sequence, ease of teaching, etc.) can be obtained during this stage to aid in providing a rational basis and avoiding major errors in the design and development of the complete program. It is further contended that long-term research and evaluation plans should begin during step 3 so that the findings of each study will add systematically to the knowledge of the students' learning and to ways of developing better educational programs. To carry out such long-term plans, development of a set of "anchor" evaluative instruments will be required. These "anchor" instruments should be independent of the specific terminology and approaches used in any version or edition of the program. They should cover the program's objectives as well as the objectives not in the program but considered important by other experts. These instruments are necessary for evaluating the objectives and effectiveness of different versions of the program and of competitive programs.

CONCEPTUAL DISCIPLINARY STRUCTURES AND THE CURRICULUM, *F. Michael Connelly,* Ontario Institute for Studies in Education

Changes in secondary school curriculums are argued for, and new materials are developed in a spirit of reaction to John Dewey and the "progressives." Two aspects of this curricular restructuring which Dewey would applaud, however, are the response of these curriculums to new disciplinary knowledge and the concern for disciplinary structure and for inquiry. It is widely held that structure and inquiry point the way to meaning and to transfer. In Dewey's terms, this amounts to the use of a new logical curricular structure.

Curriculum planners have not been clear on what they mean by "structure" and by "inquiry." By and large supported by the meaning given to structure by studies in cognitive psychology, curriculum designers have acted as if a discipline had a single structure and as if this were a matter of one or more hierarchical classifications of knowledge from facts to most inclusive concepts. According to this view, meaning becomes a matter of the properties and relations among parts of the hierarchy, and transfer becomes a matter of the analogous aspects of the most inclusive concepts of different hierarchies.

An alternative to structure conceived in terms of products of inquiry is a conception in terms of patterns of inquiry. My work in ecology is illustrative and uses as its major terms principle, problem, and fact. A pattern of inquiry is described as a process in which the subject of inquiry is bounded and analyzed in terms of a principle and is, accordingly, made ready for inquiry. For instance, one researcher may treat a school and its related units as a system of economic flows with inputs, translocations, outputs, and subsystems. Another researcher may treat the same school units as a set of structurally related parts each of which performs one or more administrative functions for the set. This illustration, furthermore, reveals the role of major classes of problems as being the source of the material terms of analysis in inquiry. In the first case, an economic problem is involved, and in the second case, an administrative problem is involved. Next, the description of a pattern of inquiry shows what kinds of data are required by the terms of the principle and problem, and how the collected data are interpreted to give a statement of outcomes embodying the terms of the principle and problem. It turns out that ecology uses four principles and attacks five problems giving twenty possible patterns of inquiry.

According to this conception of structure, meaning is a matter of the relations of fact to outcomes of inquiry. The student knows the grounds, limits, and arguments for a statement of knowledge. Furthermore, meaning is a matter of the articulation by principles and problems of the different forms of statements of knowledge arising from the different patterns of inquiry. Transfer is a matter of the degree of general use of the principles, and, in a curriculum using a method of inquiry into inquiry, it becomes a matter of intelligent habit.

SYSTEMATIC STRUCTURES FOR CURRICULUM EVALUATION, *Jeff A. Pyatte*, University of Virginia

The currently popular ideas on the structure of disciplines and their ramifications could offer curriculum theorists some badly needed assistance. Using these ideas, three levels of curriculum theory-building activity can be identified. Activities falling into the first of these levels are concerned primarily with definition, codification, and description. The second level includes those activities which are necessary in developing conceptual models for curriculum theorizing and in devising methods by which the models can be validated. The third level includes the activities which are of a judgmental or valuational character.

Activities belonging in the first and the third levels have occupied too much of curriculum theorists' time. As a consequence, the needed conceptual models and their associated methods of validation have not been sufficiently developed. Curriculum theorists must now direct their efforts toward the development of better conceptual models of curriculum and the identification of the necessary supporting models for their validation. To call attention to their methodological character, I have called the conceptual models substantive structures and the supporting models systematic structures.

The ingredients for a substantive structure have been available to the curriculum theorist for some time, and it would seem that a usable substantive structure already exists. There are also some good systematic structures which can be used to support and improve the substantive structure. I have devised and tested one such systematic structure based upon the ideas of Gagné as they apply to evaluating the effectiveness of courses of study. This model is applicable over a wider range of curriculum, and other useful models are available.

Because of the "educational revolution," there is a pressing need for more adequate models for evaluating curriculum and curriculum materials. The minimum requirements are (1) a substantive structure of curriculum into which the supporting systematic structures can be fitted, (2) systematic structures which can be experimentally tested, and (3) substantive and supporting structures which are compatible, self-modifying, and productive of new knowledge.

Session 7.2 **Division A**

CONFLICT AND INNOVATION

James M. Lipham, University of Wisconsin, Chairman

SALIENT AREAS OF CONFLICT BETWEEN ADMINISTRATION AND TEACHERS IN A MINIMALLY-INDUSTRIALIZED AREA, *John J. Seaberg, Jr.*, Southwestern Cooperative Educational Laboratory, Inc.; *Paul G. Liberty, Jr.*, Southwestern Cooperative Educational Laboratory, Inc.

This study specifies the nature of the conflict within the educational enterprise between administration and teacher organizations in three broad areas: (1) degree of importance of selected school district policies, (2) areas in which teachers should or should not be involved with policy formulation and decision-making, and (3) methods teachers should employ to obtain their goals.

A four-point Likert-type scale, consisted of 44 items, was developed to assess the three broad areas. Twenty-one administrators, 36 NEA affiliate officers, and 33 AFT affiliate officers in New Mexico school districts having both teacher organizations responded to the items. One-way analyses of variance were employed to evaluate the first two broad areas. Kendall's Coefficient of Concordance was used to show the relationship of each group's rankings for Part Three of the questionnaire.

The following results were obtained: (1) On the issue of importance of selected school district policies, the two teacher organizations were in agreement. Administrators and the teacher organizations differed significantly. (2) The two teacher organizations differed slightly on the issue of teacher involvement in areas of policy formulation and decision-making. Greatest disagreement in this area was noted between AFT affiliate officers and administration. (3) A great deal of uncertainty was found between the teacher groups and between teacher groups and administration regarding the method teachers should employ to become involved in district policy formulation and decision-making. Possible methods included (a) none, (b) informal teacher-administration contacts, (c) professional negotiation only, (d) collective bargaining only, (e) sanctions when professional negotiations have failed, and (f) strikes when collective bargaining has failed.

One notable finding was that, regarding the method teachers should use if they were to be involved in determining qualifications of applicants for the position of Superintendent, administrators preferred no teacher involvement, whereas both teacher groups preferred collective negotiations. Other findings are presented, implications discussed, and recommendations made.

A SOCIAL SYSTEM ANALYSIS OF INNOVATIONS IN SIXTEEN SCHOOL DISTRICTS, *Samuel G. Christie*, U.C.L.A.; *Jay D. Scribner*, Harvard University

The general purpose of this study was to test relationships between selected individual, structural, and system variables and the rate of adoption of educational innovations in school districts. Twenty-one hypotheses were advanced in an attempt to explain variation in adoption rates at different levels of the system. This paper focuses specifically on the variables of central interest in the study—reference group orientation, norms on innovativeness (NOI), and two system variables, expenditure and size.

Since the school-community was conceptualized as a social system, data were collected on school board members, superintendents, principals, and teachers from sixteen Southern California school districts. Scales were developed for this study to measure innovation at three levels: district, school building, and classroom. Stepwise regression was utilized to analyze the data at each level. A summary of results based on the correlations computed in this study are as follows:

1. The results of the study indicate that NOI are not significantly related to the innovativeness of teachers, principals, and superintendents. There was a significant relationship, however, between NOI and district innovativeness at the board member level. The relationship was negative rather than positive as hypothesized. This result might be accounted for in the perceptual differences of board members on district innovation. For instance, the findings showed a positive relationship between the board member's perception of district innovation and NOI.
2. A cosmopolitan reference group orientation by teachers and principals was positively related to innovation at the classroom and building level. However, this did not hold true for board members and superintendents.
3. Size of student population at both building and district level was positively related to innovation. Expenditure was significant in simple correlation; however, with years of experience of the superintendent partialled out, the relationship between expenditure and innovation drops below the five per cent level of confidence.
4. Reference group orientation seems to be a measure of individual behavior; it does not relate to policy decisions at the district level.
5. This study substantiates findings by others that size is positively related to innovation, and that expenditure is not significant when considered in isolation from other variables.

TECHNOLOGICAL CHANGE AND EDUCATIONAL OBSOLESCENCE: CHALLENGES TO ADMINISTRATORS AS MANAGERS OF CHANGE, *Richard H. P. Kraft*, Florida State University

Although a dialogue has been initiated within recent years between scientific analysts and educational planners, much of the planning work in vocational-technical education is still limited to general studies of historical data or addresses itself to the analysis of alternative programs.

The thesis has been proposed that the educational planner-administrator needs strong predictive quantitative models, suitable for testing, having cross-technology capability, and linking technology with economic feasibility. These could be used to identify long-term technological changes.

Not all administrators are willing to consider and be constrained by the requirements of the occupational end-use of their products. The problem, then, is how to develop a system or set of sub-systems which would facilitate the syndromization of occupational requirements and occupational-technical education planning objectives.

Existing automated counterparts as substitutes for human control and communication processes are discussed in relation to technical education planning. They may provide the planner-administrator with the basis upon which to build predictive instruments for future changes in occupations. The social demand approach to educational planning is emphasized by contrast with economic analyses and operations research methods.

Three central "follow-up" considerations to the objectives of a vocational-technical education planning process are: (1) to ensure that action follows policy; (2) to develop in administrators an improved understanding of technological developments, including automation; and (3) to give primary attention to those generic-system environment relationships which are affected by, and which in turn affect, planning decisions.

To the extent that recent technological developments emphasized the need for long-range planning, a systems look as outlined in this paper may provide a methodological basis for interdisciplinary, planning-oriented research. Work in progress at the Educational Systems Development Center at the Florida State University, dealing with the social demand approach to educational planning, is described in reference to the specific considerations already discussed, and the paper concludes with comments on the possibilities and limitations of vocational-technical education planning and its integration in a broader framework of social planning.

A TEST OF A CONFLICT THEORY OF ORGANIZATION, *Robert B. Brumbaugh,* West Virginia University

Conflict theories of organization imply potential exploitation of the individual by the formal organization. While proposing generally that the individual may be victimized, few such theories have addressed themselves to specific formulations of the phenomenon. This paper reports an empirical test of a theory focusing on one of the explicit ways the school as an organization may be exploitive of the individual—exploitive in the sense that only certain of the individual's interpersonal needs may be salient or relevant to the hierarchically-contrived authority system.

We hypothesized that, of the three interpersonal needs measured by Schutz's FIRO-B Scale—inclusion, control, and affection—(1) control will be the most salient need perceived in subordinates by immediate superiors, and (2) the immediate superior of a given subordinate is more likely to evidence a control orientation toward that subordinate than is a superior further removed hierarchically (control orientation is operationally defined as the degree of accuracy achieved in predicting the FIRO-B control needs of another). The FIRO-B predictions of 29 elementary and secondary school principals for one randomly selected teacher from each principal's school were compared with the FIRO-B predictions for these same teachers made by upper-echelon school district central office professional staff members. A large West Virginia school district provided the population for the study sample.

Among other findings, chi-square analyses of the data revealed that (1) the FIRO-B control needs of the teachers tended to be more accurately predicted by the principals than were the inclusion or affection needs, but that (2) the principals were no more likely to be accurate in predicting the teachers' control needs than were the central office professional staff. In predicting each other's FIRO-B needs, however, principals tended to be more accurate in predicting the teachers' control needs than was conversely the case. Implications are drawn for theory and research.

Session 7.3

Division A

POLITICAL SOCIALIZATION AND OTHER APPROACHES TO THE STUDY OF SCHOOL OUTCOMES

Conrad Briner, Claremont Graduate School, Chairman

CONTEXTUAL EFFECTS ON EDUCATIONAL ASPIRATIONS: SOME CANADIAN EVIDENCE, *Emil J. Haller*, Cornell University; *Barry D. Anderson*, Washington University (St. Louis)

The purpose of this research was to assess the relative influence of certain contextual variables on educational aspirations. Following the publication of the Coleman report, *Equality of Educational Opportunity*, the contextual effects phenomenon received increasing attention from educational researchers and administrators. The Coleman report is, however, only one of many studies which indicate that the social context of a school has an independent influence on students' academic behavior. In these studies, though, the relative strength attributed to contextual variables has ranged widely. More important, with rare exceptions, these studics have not been longitudinal, nor have students actually been subjected to varying contexts in experimental or quasi-experimental designs. This research was an attempt to rectify some of these problems.

A cohort of all secondary school pupils in Ontario was followed from grade 9 to grade 12. All pupils who transferred to a second school at the end of the 9th grade were selected from this cohort. The social context of both sending and receiving schools was indexed using socioeconomic variables. Educational aspirations were determined for each pupil in the 9th grade and again in the 12th grade. Changes in aspirations between these two grades were then regressed on differences in school contexts, with the individual characteristics of ability, family socioeconomic status, and sex controlled in a multiple regression analysis. Results indicate that context accounts for little of the variance in aspirations.

This study suggests that administrative attempts to "homogenize" social classes in secondary schools (e.g., through busing) in order to influence educational aspirations may be less effective than has been supposed.

COMMUNITY CHARACTERISTICS RELATED TO PERCEIVED STUDENT SATISFACTION AND ACHIEVEMENT IN 100 PUBLIC JUNIOR COLLEGES, *Vernon L. Hendrix*, Dallas County Junior College District

A recently completed study identified 17 variables (Battery A) descriptive of the major dimensions of 100 communities in which public junior colleges were located. Thirteen of these variables described the social, cultural, economic, and demographic characteristics of the communities. Four of them described the environmental press preferences of the faculty and students. Using the Junior College Environment Scales, 4 dimensions (Battery B) describing the environment of the colleges were discovered. The primary analyses performed in this study concerned the relationships between these two batteries and 35 items (Battery C) which enabled students to indicate their satisfaction with college and rate their achievement of certain educational goals. This paper only reports the relationships between Battery A (Community Characteristics) and Battery C (Satisfaction and Achievement), using canonical analysis.

The results indicate that there are inherent conflicts. Certain satisfactions and perceived achievements are apparently less likely to occur in the presence of certain community characteristics. This implies that there are limitations imposed by the nature of the community which prohibit the maximization of certain student satisfactions and achievements by deliberate manipulations of the college environment.

POLITICAL SYSTEMS, PUBLIC SCHOOLS AND POLITICAL SOCIALIZATION, *George R. LaNoue,* Teachers College, Columbia University; *Norman Adler,* Teachers College, Columbia University

From the initiation of the public school system in the United States, it has been assumed that the schools were essential to a healthy democracy. The school's contribution of skills is obvious, but the nature and effect of school-created political values is not so apparent. It is the purpose of this paper to point out some of the linkages between the political system, the public schools, and the political socialization process, and to indicate how these linkages can be studied.

Most descriptions in the education literature on the educational policy process depends on the nebulous "establishment" concept, and until recently political scientists have rarely studied this particular area. Consequently the authors have sought to utilize David Easton's systems approach by creating a systems model of the educational policy-making process. When the political inputs into the school are examined, the authors find that the traditional inputs were basically system supporting, while the more recent inputs into the policy-making process are system modifying, bringing in new groups whose demands and attitudes are more heterogenous. The lack of political research on the school makes conclusions about the role of the converter in the model uncertain, but the clues suggest that the organization and ethos of the school lead toward conformist values and the status quo. Since lower class children receive most of their political values from schools, the public schools bear a large share of the responsibility for the general lower class failure to support civil liberties. Middle class youngsters are more influenced by family norms, are less arbitrarily treated in the schools, and are more likely to attend college. Consequently they are better able to transcend public school political socialization.

In summary, then, the public schools do quite an adequate job in encouraging respect for regime and stability, but their ability to promote a tolerance for diversity and for civil liberties is not so satisfactory.

PSYCHOPHYSIOLOGICAL CORRELATES OF FEMALE TEACHER BEHAVIOR AND ORGANIZATIONAL OUTPUTS, *Carolyn Lipton Ellner,* Claremont Graduate School

The purposes of this study were (1) to investigate the relation between measures of individual differences in the functioning of the Autonomic Nervous System (ANS) to indices of teacher performance, emotional behavior, and attitudes and (2) to discover whether congruence of teacher role expectations between teacher and principal bore a relation to (a) measures of teacher effectiveness and (b) indices of stress, satisfaction, and goal achievement.

A standardized inventory and other instruments developed especially for this study were administered in sequence personally to each of sixty-three female teachers for whom estimates of autonomic balance had been previously obtained and to each of their principals. Measures of emotional behavior, teacher role preference and role perception, principals' teacher role expectations and perception of teacher role enactment, assessments of effectiveness, organizational attitudes, and stress were obtained.

A systematic association was found between original estimates of autonomic balance (A Scores) and self-assessments of emotional behavior, educational goal preferences, organizational behavior, and role conflict. Personality factors were significantly correlated with teacher role behavior and goal emphasis, with accurate interpretation of principals' role expectations, and with teaching satisfaction and perceived stress.

Rated effectiveness appeared to be associated with congruence between principals' role expectations and their perception of teachers' role enactment and with a preference by the teacher to cope with educational problems within the classroom.

The study concluded that (1) estimates of autonomic balance are related to personality and teacher role behavior; (2) emotional behavior is associated with teacher role behavior; (3) teacher and principal appear unable to report the others' normative standards for teaching behavior and that lack of communication in the school exists in this regard; (4) relative goal emphasis has no relation to rated effectiveness; (5) rated effectiveness is the product of congruence between principals' role expectations and their perception of teachers' role enactment; and (6) positive organizational sentiments are related to the principals' assessments of teacher behavior and to personality factors.

Session 7.4 **Division B**

INSTRUCTIONAL MATERIALS

Robert Bridgham, Stanford University, Chairman

THE RELATIONSHIP BETWEEN STUDENT INTEREST AND STUDENT LEARNING WITH NEW CURRICULAR MATERIALS,

F. Lincoln Grahlfs, Sociological Resources for the Social Studies

SRSS has tried, among other things, to produce materials which interest students. However, there are segments of any curriculum which, though vital, are not inherently interesting to students. How far should a project go in considering student interest to produce effective learning?

Between January, 1967 and June, 1968, twelve two-week units in sociology were taught for evaluation. Each was taught to several hundred students in various parts of the United States.

Student learning was assessed by the use of a 35-item multiple-choice test administered to control and stimulus groups. After the unit had been completed, both students and teachers were asked how interesting they found the material.

On the test for each of the twelve units, the stimulus group had a raw score mean which was significantly higher than that of the control group. In each case, also, there was a high positive correlation between ability and test scores. The correlation between student interest and test scores, however, was low or negative for many episodes.

To be sure, we want always to make materials as interesting as possible. The implication of these findings, though, is that we should not be unduly concerned about interest. Curriculum writers cannot afford the luxury of a "popularity contest." If, for example, students are to be given a genuine appreciation for social sciences, they must be exposed to charts, tables, and statistics. Many social studies students, and teachers, too, have a negative set toward these things. In spite of this, the test results show that a large proportion will assimilate them.

THE RELATIONSHIP BETWEEN STUDENT PERFORMANCE ON NEW CURRICULAR MATERIALS IN SOCIOLOGY AND TEACHER PREPARATION IN SOCIOLOGY, *William M. Hering, Jr.*, Sociological Resources for the Social Studies

As curriculum projects release increasing amounts of materials for classroom use, concern for the teacher's ability to make the best use of that material has been expressed. Projects whose subject matter is seldom emphasized in teacher preparation requirements have been especially concerned. Sociology is one example. Efforts have been made to encourage more and more summer and in-service institutes, local workshops, and so on. All of this activity has been based on the assumption that materials will be more effective if teachers are well prepared in the subject.

Sociological Resources for the Social Studies has data from the field trials of twelve two-week units in sociology. Student performance was measured in terms of gain from pre- to post-tests. Ss were approximately 8,000 students. The sample was categorized according to ability level as measured by a standardized verbal reasoning test. Over 250 teachers were involved in the trials, representing a wide range of educational experience and preparation in sociology.

Student gain, controlled by the verbal ability measure, was correlated with teacher preparation in sociology as indicated on a teacher questionnaire. Results indicate that the projects may be assuming too much. There was no significant correlation between student performance and teacher preparation; in fact, there was some trend to the contrary. Teachers with little preparation in sociology did show significant gain in student performance.

The reasons for this finding are not clear. It may be that the sociology presented by the project differed markedly from that learned by the teachers, or that the methodology implicit in the project materials was unfamiliar to the teachers. Possible implications are that little special training is necessary to teach project materials, or that something more than additional study in subject matter is needed to implement new curricular materials. Future evaluation will include teachers who have participated in institutes which focused on project materials, and this additional factor will be included in later research.

COMPONENTIAL EVALUATION OF CREATIVITY INSTRUCTIONAL MATERIALS, *Susan J. Bahlke* and *John F. Feldhusen*, Purdue University

The purpose of this research was to evaluate the componential training effects of a series of 28 programs designed to develop children's abilities in the divergent thinking functions of originality, fluency, flexibility, and elaboration.

From a population of 4th, 5th, and 6th grades in an urban school system, 54 classes, 18 at each level, were randomly selected and assigned to 9 treatment conditions. Each program consisted of a presentation (P) of some principle for improving creative thinking, a story (S), and a set of printed exercises (E). Each of the three parts was presented separately, in pairs, and all together for a total of seven experimental arrangements. In addition, 2 control groups were used.

The materials were administered twice a week for 14 weeks by the classroom teacher. The *Torrance Tests of Creative Thinking* were administered as the pretest and posttest.

A componential analysis using a five-day analysis of variance for three treatment conditions (P, S, and E, each 2 levels), sex, and grade was used. Significant F's were obtained from P in verbal originality ($p < .05$) and from E in nonverbal fluency ($p < .01$), nonverbal flexibility ($p < .01$), and verbal originality ($p < .01$). Interactions between grade and all three treatment conditions were present in verbal flexibility ($p < .01$) and verbal originality ($p < .01$). Interactions between various treatment conditions were also present. P x E yielded a significant F in verbal originality ($p < .01$), and S x E yielded significant F's ($p < .01$) for all variables except nonverbal originality.

Thus, three types of creativity training procedures produced differential gains in creative thinking when used singly and in various combinations. The results suggest that components of instructional packages should be evaluated separately to determine which are producing the training effects.

NEGRO ATTITUDES TOWARD PICTURES FOR JUNIOR HIGH SCHOOL SOCIAL STUDIES TEXTBOOKS, *Nicholis Poulos*, Detroit Public Schools

The purpose of this investigation was to determine the attitudes of Negro members of parent-teacher organizations toward hypothetical pictures of Negro personalities and events for junior high school social studies textbooks. The need for the study was based on the general uncertainty currently prevalent among large-city boards of education, educators, and textbook publishers regarding Negro-oriented content in school textbooks.

Members of parent-teacher organizations of elementary and junior high schools in the metropolitan area of Detroit, Michigan, served as the respondents. Fifty-five pictures portraying Negro historical personalities and events were selected, made into picture slides, and projected on a screen for group viewing. The respondents viewed the pictures and recorded their attitudes on a five-point scale from "strongly approve" to "strongly disapprove."

It was found that:

1. Males were generally more favorable than females to the pictures.
2. Pictorial approval was in proportion to the increasing occupational status of the respondents.
3. The pictures were most favorably received by members of community organizations, somewhat less favorably by community leaders, and least favorably by nonparticipants.
4. Respondents possessing at least one year of college education were more favorable to the pictures than those with less education.

There were no significant differences in response among the factors of (1) age, (2) home stability, and (3) number of children in the family attending school.

The respondents were more receptive to the pictures depicting eminent individuals who enhanced the Negro image rather than to those which portrayed the hardships and privations endured by the race during the past or present. The illustrations of the twentieth-century period were generally most favorably received, whereas those of the slavery era were least favorably received. Certain factors in the backgrounds of the respondents had a significant bearing on attitudes toward the composite pictures and also toward those classified in historical categories.

Session 7.5 **Division C**

PROGRAMMED INSTRUCTION

William Loadman, Michigan State University, Chairman

AN APPLICATION OF GUILFORD'S STRUCTURE OF INTELLECT TO PROGRAMMED INSTRUCTION, *Robert L. Baker,* Southwest Regional Laboratory for Educational Research and Development

If Guilford's cubical model of the intellect has any potential for the preparation of learning materials, then identification of the relevant learner characteristics and matching those characteristics with appropriate instructional stimulus materials should optimize the efficiency of self-instructional programing procedures. It should also permit the generalization of a psychometric model to the solution of specified instructional problems.

The objectives of the project were to determine: (1) the comparative effectiveness of symbolic and semantic content-based linear programs for teaching the rules of logic to eighth-grade students, (2) the relationship between intellectual abilities as measured by selected tests available in Guilford's compendium of aptitude tests and achievement in each of the two program variations, and (3) the efficiency of matching pupils with the instructional program variation on which they have the highest predicted success.

The general research strategy followed was to develop two variations of a self-instructional program on the basis of cues provided by the semantic and symbolic aspects of the content dimension of Guilford's model.

A series of three studies was conducted with the following results:

1. Random assignment of 160 Ss to the two program variations

yielded almost identical criterion mean scores for the semantic and symbolic forms.

2. Differential performance on the two forms of the logic program cannot be explained in terms of the content dimension of Guilford's model, since in both cases the semantic tests contribute the overwhelming proportion of the total and explained variances.

3. Guilford tests loaded on factors from the *operations* and *products* dimensions had greater utility for predicting performance on the two logic program variations.

While use of Guilford's model to aid in the prediction of differential performance on instructional tasks has considerable potential, use of his model to suggest design aspects of the alternatives themselves, either the description of instructional objectives or the specifications of instruction, is not supported by these data.

A COMPARISON OF VERBAL STATEMENT, SYMBOLIC NOTATION, AND FIGURAL REPRESENTATION OF PROGRAMED STRUCTURAL GRAMMAR, *Wayne C. Fredrick*, Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin

Three versions of programed lessons in structural grammar were developed to teach concepts such as basic sentence patterns, sentence parts, and methods of determining parts of speech (affixes, position, markers, test-sentences). In one version all concepts were presented as verbal statements, i.e., "A noun phrase consists of noun markers plus a noun." A second version, after initial verbal presentation, used symbols i.e., $NP \rightarrow NM + N$, while a third version used the symbols arranged spatially in sentence trees. The content was held constant by maintaining a frame-by-frame similarity among the three versions with only the particular mode of representation varying.

Three eighth-grade classes (72 *S*s) studied the lessons as part of their regular English curriculum for one week. Each version was in use simultaneously in each class along with control *S*s who studied programmed lessons in poetry interpretation. Posttests given one day and 14 days after the treatment showed that all three groups learned and retained more grammar than the control group. The symbolic and figural versions were superior to the verbal presentation, and an interaction between ability and mode of representation was present. This interaction showed that only low ability *S*s did not benefit from the figural version, while only high ability *S*s mastered the verbal presentation. The symbolic notation produced significant learning at all ability levels. The teaching of verbal concepts seems to be facilitated by appropriate use of symbols and diagrams, provided the *S*s understand these non-verbal materials.

THE EFFECT OF CREATIVITY, RESPONSE MODE, AND SUBJECT MATTER FAMILIARITY ON ACHIEVEMENT IN PROGRAMMED INSTRUCTION, *Sigmund Tobias*, City College, City University of New York

This study consisted of a 2x2x3 ANOVA design. The first classification comprised two levels of creativity determined by scores on the Remote Associates Test. The second level consisted of two response modes: (1) constructed response, and (2) reading the material cast in the form of completed sentences. The third level was made up of three measures of *S*'s (*N* = 100) achievement on familiar material, technical material requiring a verbal response, and technical material requiring a pictorial response. The technical part of the instructional material consisted of a revised version of a linear program dealing with the diagnosis of myocardial infraction. The familiar material consisted of 54 introductory frames dealing with the incidence and risk of contracting heart disease. Blackout ratios for both the technical and the familiar material were low.

Previous research and theory led to the prediction of no main effect, although interaction between creativity and novelty of subject matter, and between response mode and novelty of subject matter, were expected. The results indicated that, contrary to prediction, the achievement of the high creativity group was significantly above that of the low group on all subject matter tested. The predicted interaction between response mode and familiarity of subject matter was obtained. There were no significant differences between the constructed response and reading group on the familiar material, although these groups did differ on technical material requiring verbal response ($p > .05$) and on material requiring a pictorial response ($p = > .01$).

These findings confirm, in one investigation, that discrepant findings in prior research on response mode may well be attributed to the degree to which the subject matter was familiar to the Ss. Furthermore, the data suggest that blackout ratio of a program may be less important than the familiarity of the subject matter in predicting which response mode is likely to lead to the highest achievement.

THE EFFECTS OF PROGRAMMED INSTRUCTION IN PRODUCTIVE THINKING ON CREATIVE PROBLEM SOLVING ABILITIES AND ATTITUDES AMONG PUPILS IN GRADES FOUR THROUGH SEVEN, *Donald J. Treffinger*, Purdue University; *Richard E. Ripple*, Cornell University

The effectiveness of the *Productive Thinking Program* (Covington, Crutchfield, and Davies, 1966) in developing creative thinking and problem solving abilities and related attitudes among 370 pupils in grades four through seven was examined.

At all four grade levels studied, instructed pupils (N = 184) attained significantly greater scores than control pupils (N = 186) on Covington's (1967) *Childhood Attitude Inventory for Problem Solving*. Such differences were found for general attitudes about creative thinking and problem solving and for total test score, but not for expressions of self-confidence about creativity and problem solving.

There were no significant indications of transfer from the programed instructional materials to several measures of creative thinking and problem solving at any of the four grade levels studied. Criteria included: ten scores derived from the Torrance Tests of Creative Thinking (Verbal Form A, 1966), Getzels and Jackson's Make Up Problems Test (1962), a general problem solving test, an Arithmetic Puzzles test, and an Arithmetic Problem Solving test which represented text book number problems.

Results were interpreted in terms of three general factors: the severe conditions under which the program was administered, criterion difficulty and degree of similarity to the instructional material, and the complexity of creative thinking and problem solving abilities.

Session 7.6 **Division C**

PSYCHOLINGUISTICS NO. 2

Loren S. Barritt, University of Michigan

CHILDREN'S CONCEPTION OF WORD BOUNDARIES AS A FUNCTION OF DIFFERENT LINGUISTIC CONTEXTS, *Marjorie H. Holden* and *Walter H. MacGinitie*, Teachers College, Columbia University

The purposes of this study were to investigate (1) young children's understanding of word boundaries in the auditory mode, and (2) the relationships between this understanding and standard predictors of later reading achievement. Ss were 84 pre-literate kindergarten children from a middle-class community.

Ss were individually tested. They listened to a taped utterance until they could repeat it correctly. Then they tapped on a row of markers—one tap for each word—while repeating the utterance. Each child was tested on 12 to 20 utterances. Each utterance was presented to at least 24 children.

The present study revealed that: (1) there was a strong tendency to compound function words with content words, especially with the content words directly following; (2) certain free forms correctly isolated by most Ss in one context were compounded with other forms when they were used in a different construction; (3) certain compoundings that violated printing conventions were correct according to linguistic definitions discussed by Greenberg (1957). Four children tested at the end of first grade scored significantly higher than the kindergarten children, as did a small sample of older, severely retarded readers.

These findings support earlier conclusions that auditory awareness of word units begins with content words. The ability to isolate free forms in an utterance depends also, however, on (1) sentence type, (2) function of the word within the sentence, and (3) position of word in relation to other words in the sentence.

The *Gates-MacGinitie Readiness Skills* test will be given to these same Ss at the beginning of first grade, and correlations between readiness variables and the ability to identify free forms will also be reported and analyzed.

PATTERNS OF EFFECTIVE CONTEXTUAL CONSTRAINT IN SENTENCES, *Lewis W. Pike*, Educational Testing Service

An important feature of language is redundancy, which helps preserve the essence of a message even when an occasional word is missed.

One index of redundancy is the percentage of subjects who can correctly replace an omitted word. Studies using this measure have shown that, on the average, contextual constraint operates over very short distances, and that its effectiveness drops off sharply as distance increases. The present study investigated structural relationships and individual constraint effects underlying these distance-constraint findings.

Constraint associated with selected structural relationships was determined for 81 sentences, all having the same structure, by means of one- and two-word omission patterns. The constraint imposed by a subject noun on a verb, for example, was found by comparing verb predictability when full context was available to that when both the verb and subject noun had been removed. Analyses were based on 81 responses from each of 216 subjects.

The distance-constraint curve reported in previous studies was characteristic neither of individual sentences, nor of data averaged over all 81 sentences, each having the same structure. The distribution of individual constraint effects yielded two surprises. Many individual effects were large, despite a small overall mean effect. Moreover, removing a context word often *increased* test-word predictability, sometimes dramatically.

Small mean differences in predictability associated with the effect of distant context have consistently been interpreted as indicating that distant effects are necessarily rare. The discovery of frequent reverse effects casts doubt on such an interpretation, and re-opens the question of long-range constraint.

Such information will be useful in understanding how redundancy functions in the perception of spoken and written language.

RECALL OF THE BASE AND SURFACE STRUCTURE FORMS OF SENTENCES, *Diane Simison*, Indiana University

In transformational grammars, the base structure of the sentence, as reflected in the Base P marker, reflects the formal specification of meaning. Sentence surface structure may be different in form from the base P marker, so that a sentence with the base P marker SVO+SVO may have five surface structures: SVO+VO, SVO+O, S(V+V)O, S+SVO, SV+SVO.

If the base form is more psychologically real to a speaker, he will make transformations of the surface structure during sentence processing. The hypothesis is that a speaker, on hearing a sentence, transforms it to its base structure. These transformation processes will take more room in memory for sentences of the surface structure type. A speaker will have more room in memory after processing sentences of the base P marker form, and will make more errors on recall of surface structure than on base structure.

The criterion used to measure room remaining in memory after sentence processing was number of words recalled from a word list presented after presentation of the sentence. 30 subjects were divided into two groups. Group I received sentences in base form that group II received in surface form. Within subjects, each S received ten sentences of each of five types: five in base P marker form, and five in surface structure form. After each sentence was heard, the subject heard a list of eight words. Word lists were counterbalanced across sentence types.

A Lindquist Type VI mixed design was used with mean number of words recalled and number of errors on sentences as dependent variables. No difference was found in mean number of words recalled either for structure or for type of sentence. However, subjects make significantly more errors on recall of the surface form of the sentence and on recall of SVO+VO and SVO+O sentence types.

PHONETIC SYMBOLISM IN ADULT NATIVE SPEAKERS OF ENGLISH: A REPLICATION, *Robert D. Tarte* and *Loren S. Barritt*, Center for Research on Language and Language Behavior, U. of Michigan

The evidence for or against the phenomenon of phonetic symbolism has been inconclusive. Tarte and Barritt (1968) have demonstrated that there are unrecognized, consistent bases which English speakers use to respond linguistically to geometric figures. The present study is a replication of this previous work, with adult native speakers of French.

A forced-choice, paired-comparison test containing eight geometric figures paired with one another and presented with nine monosyllabic sounds was presented to *Ss*. There were 252 trials. *Ss* were asked to decide which of the two figures presented on each trial was most like, or went best with, the monosyllabic sound presented on audio tape. The nine sounds used were "was," "wus," "wis," "das," "dus," "dis," "kas," "kus," "kis."

Results from the present study confirmed the conclusion from earlier studies with English-speaking adults that *Ss* respond in a consistent manner when asked to label geometric figures. With English-speaking *Ss* the size dimension exerts the strongest effect. However, with native speakers of French, consistent responses are produced across the shape dimension rather than the size dimension.

Results from these studies support the conclusion that phonetic symbolism is a culturally determined, rather than universal, phenomenon.

LANGUAGE STRUCTURE AND THE FREE RECALL OF VERBAL MESSAGES BY CHILDREN, *Paul Weener*, The Pennsylvania State University

This research investigated the development of the ability of children to utilize the between-word structure of language, independent of inflectional and intonational cues. An instrument was designed to measure the effects of contextual constraint and syntax on the recallability of verbal messages. Four task levels were defined by all possible combinations of high and low contextual constraint and high and low syntax. Five items, each made up of two five-word sentences, constituted the stimulus material at a given task level. The words used at different task levels were matched for frequency and form class. Twenty-four normal *Ss* from kindergarten, first, second, and third grades were tested at all four task levels. The items were presented as a free recall task at the rate of one word per second with minimal intonation and inflection. Responses were scored for total number of words recalled. It was hypothesized that the effects of each of the two structural factors on recall would be significant, and that both factors would interact with grade level, with older children obtaining more benefit from the structured messages.

The main effects of age and contextual constraint were significant, but no significant differences were observed between the high and low syntax conditions at any grade level. Contextual constraint interacted with grade level in the predicted manner with the difference between high and low contextual constraint increasing with grade level. No interaction between syntax and grade level was observed. The lack of an effect due to syntax implies that children rely on the intonational or inflectional features of a sentence to provide cues regarding the structure of sentences, and that without these suprasegmental features the structure of the sentence is not utilized.

Session 7.7 **Division C**

TEACHER BEHAVIOR

Chairman: To Be Announced

A COMPARISON OF THE ROLE OF TEACHERS IN CAI CLASSROOMS AND TEACHERS IN TRADITIONAL CLASSROOMS, *Russell A. Hill* and *Norma F. Furst*, Temple University

This study compared the role of teachers working in a Computer-Assisted Instruction (CAI) classroom with the role of the teacher working in a traditional classroom. The study of differences included (1) teacher behavior data, (2) teacher perception data, and (3) observer reports. Ten secondary schoolteachers were studied. All the teachers were teaching the same subject at the same grade level both in a CAI classroom and in a traditional classroom within the same school with the same type of students.

Three types of data were collected about each teacher. Records were made of teacher classroom behavior using a systematic observational instrument. Perceptions of the teachers toward their roles in the different classrooms were recorded through the use of a structured questionnaire and an unstructured interview. The structured questions focused upon the planning operations, the actual execution of classroom procedures and evaluative procedures of the teacher as they worked in the two different types of instructional modes. Observer anecdotal records provided further descriptive evidence of differences.

The systematic observation behavior data for the teachers in CAI and the same teachers in traditional classrooms were treated by analysis of variance using repeated measures. The questionnaire data and the anecdotal descriptive reports were summarized and related to the quantitative data.

Significant differences were found in the behavior of the teachers in the two different types of classrooms. Differences were also found in the perception of carrying out the tasks. These differences were further amplified by the observer reports. In general, (1) the subject oriented informing activity of the teacher in a CAI classroom was sharply reduced, and (2) the teacher in the CAI classroom tended to relate primarily to individuals and fragments of the class rather than to the total class much more often than the teacher in the traditional classroom.

AN EXPERIMENTAL INVESTIGATION OF THE EFFECTS OF CONCEPTUAL VAGUENESS ON SPEAKING BEHAVIOR, *Jack H. Hiller*, Office of Personnel Operations, Department of the Army

Empirical measures tied to a psychological construct termed *vagueness* were found to bear statistically significant negative correlations with measures of student writing and teacher lecturing effectiveness. The experiment reported here was designed to test the hypotheses that the negative correlation in the teacher study reflected relatively poor knowledge and insufficient preparation. Each of 24 male subjects, who acted as teacher lecturers for the experiment, received two treatments before lecturing: (1) they were given a High Knowledge of a Low Knowledge lesson, and (2) they were given time to prepare, High Preparation, or only a few seconds, Low Preparation. As predicted, Ss receiving the HK lessons exhibited a significantly lower proportion of vagueness responses in their lectures than the LK group ($p < .01$). The Preparation manipulation did not independently affect the level of vagueness responding, but did demonstrate a tendency toward interaction with Knowledge ($p \leq .10$). The efficacy of the experimental treatments was further demonstrated by the finding that HK subjects earned higher scores on tests for comprehension of the manipulated lesson material. As an alternative, but not contradicting hypothesis, it was suggested that anxiety and lowered self-confidence may have acted to mediate the increase in vagueness responding.

THE RELATIONSHIP BETWEEN TEACHER VOCABULARY USAGE AND THE VOCABULARY OF KINDERGARTEN AND FIRST GRADE STUDENTS, *Robert E. Jester*, Institute for Development of Human Resources, University of Florida

The purpose of this study was to investigate the relationship between the words used by teachers at the kindergarten and first grade level and the children's knowledge of words. Twenty-four kindergarten and first grade teachers from Alachua County near Gainesville, Florida, were selected on the basis of race and socioeconomic background. Tape recordings were made for approximately one hour per teacher at randomly selected intervals during a normal classroom day. Word lists were then prepared from the tape recordings and the children were tested for their knowledge of words on the lists.

An hypothesis was made that teachers would generally use words beyond the comprehension level of their students. This hypothesis was based on comments made by supervisors of elementary education teachers and by a systematic search of the research literature which indicates that there is a "language gap" in materials prepared for young children and the children's sophistication in language usage. In addition it was expected that effects would be found due to differences in socioeconomic background and grade level.

The study was cast in the form of a five-way factorial design including independent variables of race, socioeconomic background, and grade level for both the teachers and the pupils. The dependent variable used to test the basic hypothesis was a congruence score based on the children's knowledge of the words their teachers used.

Preliminary data analysis indicates that the basic hypothesis is not supported and that there is a definite trend in the opposite direction. That is, the children are able to comprehend words at a more advanced level than the words used by the teachers suggest. Results will be discussed in terms of the independent variables and implications for classroom practices and future research will be considered in terms of those results.

MULTI-VARIATE STUDY OF TEACHER BEHAVIOR AND PUPIL ACHIEVEMENT, *Jack Sloan*, University of Wisconsin—Milwaukee

This study investigated the interrelationships among optimally discriminating linear composites of instructional unit test item variables and observed instructional behavior differences among the teachers of the unit of instruction. The study deals with the learning of a unit on arithmetic operations with rational numbers named as decimals. Subjects were 1083 sixth grade pupils and their 43 teachers.

The pupils were given two tests: a pre-instruction test on mathematical operations, and a post-instruction test on the unit on decimal-numbers. After instruction, but before the post-instructional test, each teacher was asked to sort the post-instruction test items into six categories, such that each category contained items that demanded the same skills and knowledge. Next, each teacher was asked to sequence the objectives in the same order used when the unit was taught. Finally, each teacher was asked to rate each objective on a Likert scale according to the amount of emphasis the teacher had placed on the objectives when the unit was taught.

Procrustes transformation was used to generate a scalar index of the least-squares fit between each teacher's post-test item sorting categories and the factor pattern matrix generated from the responses of the pupils to the post-instruction test. Scores of each teacher on the factors of the matrices generated in the item sequencing and emphasis rating tasks were computed, as well as class mean pre-instruction and post-instruction test scores. Demographic data on teachers and classes were also gathered.

Discriminate function analysis was employed to generate linear composites of the post-instruction test item variables. Correlation analysis was used to investigate the interrelationships among the various linear composites of test item variables, the measures of teacher behavior, and the demographic measures.

VERBAL TEACHING BEHAVIOURS AND THEIR RELATIONSHIP TO PUPIL ACHIEVEMENT, *Clifford J. Wright,* Christchurch Teachers' College, New Zealand

The purpose of the investigation was to determine which of certain selected teaching behaviours correlated significantly with a measure of pupil achievement. In addition, the study was designed (i) to provide a description of the teaching tactics of a group of New Zealand teachers, and (ii) to indicate differences among three selected groups of teachers.

The study involved 17 teachers of Grade 3 New Zealand children in 51 lessons on a nature science unit, the content of which was carefully controlled. Six of the teachers were experienced "superior" teachers and six were teachers' college students on their first teaching practice assignments. The third group were teachers' college students who had had a course in the analysis of teaching, followed by a micro-teaching course during one of their practice-teaching sections.

The lessons were tape-recorded and transcribed. The tape-scripts were analyzed to provide 61 variables of teacher behaviour. These were didactic in nature, falling for the most part into categories corresponding to three of the pedagogical moves identified by Bellack and Davitz (1963). Among the structuring moves identified were: pre-solicitation structuring; post-solicitation structuring; and summary statement closure. Solicitating moves included: single and multiple soliciting; open and closed soliciting; redirected soliciting; and six types of reciprocated soliciting (seeking amplification, extension, clarification, justification, verification or repetiton of pupil response). Teacher reaction moves included: simple and complex comment; use of reward and praise; managerial and urging comment; the use of continuant soliciting; and a number of kinds of reflection or repetition of pupil response.

The criterion measures of pupil achievement were scores on an achievement test corrected for pupil intelligence and prior knowledge of nature science concepts.

Significant relationships between certain teaching behaviours and pupil intelligence, prior knowledge and achievement are reported.

Session 7.9 **Division D**

STUDIES OF ACHIEVEMENT AND INTELLIGENCE

Samuel D. McClelland, New York Public Schools

PATTERNS OF ACADEMIC ACHIEVEMENT IN GRADE 12 CHEMISTRY AND THEIR RELATIONSHIP TO PERSONAL, ATTITUDINAL AND ENVIRONMENTAL FACTORS, *Alexander Even,* The Ontario Institute for Studies in Education

Patterns were sought in groups of profiles of residual achievement scores in the processes of Knowledge, Comprehension, Application, and Analysis as described in the *Taxonomy of Educational Objectives: Cognitive Domain.* A sixty-item end-of-course test, administered to 2339 Grade 12 General Course students in thirty randomly-selected Ontario schools, served as the criterion instrument.

Independent variables contributing to the explainable variance of the achievement scores were selected by the Automatic Interaction Detector program; these variables were, in order of decreasing importance, mathematics aptitude, verbal aptitude, immediate educational plans, and the Theoretic-Immediate and Prudent-Theoretic scores of Edwards' and Wilson's Inventory of Choices. All other variables studied made unimportant contributions to the variance of the chemistry scores. A prominent mathematics aptitude—by—verbal aptitude interaction was observed.

Congruent or parallel patterns involving all four *Taxonomy* categories studied did not appear in any groups of students. When patterns were confined to Categories 2.00, 3.00, and 4.00, congruent patterns appeared in all underachievers and in all overachievers; parallel patterns appeared in groups isolated by the AID program on the basis of immediate educational plans. Normal achievers did not exhibit congruent patterns in any grouping.

The underachiever pattern is characterized by no significant difference in score between Categories 2.00 and 3.00; a significant drop to Category 4.00 is evident. The overachiever pattern has a significant peak in Category 3.00; the drop to Category 4.00 is twice that observed for underachievers. College-bound students showed a small but significant increase in pattern level over that of non college-bound students, but both patterns were of the same shape and were not distinctive.

Anomalous properties observed for Category 4.00 may be due to insensitivity of the criterion measure or to a measure of ambiguity inherent in the structure of the higher levels of the *Taxonomy.*

SOME DIMENSIONS OF CREATIVE THINKING, SCHOOL ACHIEVEMENT, AND INTELLIGENCE IN FIRST GRADE, *Nicholas C. Aliotti* and *William E. Blanton*, University of Georgia

The present investigation was an attempt to explore the differences among figural and verbal creative thinking abilities, language and non-language intelligence, and school achievement. Specifically, two questions were posed: (1) What common variance is shared between language and non-language intelligence and figural and verbal creative thinking abilities? (2) To what extent are these measures related to school achievement?

Eighty-three first grade subjects were administered the *Torrance Tests of Creative Thinking, Metropolitan Achievement Tests, Metropolitan Readiness Tests, California Test of Mental Maturity,* and the Picture Interpretation Test (Torrance and Grossman, 1968). A total of twenty measures was obtained and intercorrelated, and the resulting matrix was subjected to a principal-component factor analysis.

Two criteria were employed to determine the appropriate number of factors to be retained for rotation. Kaiser's (1960) criterion of roots greater than one indicated five factors, while Cattell's (1966) Scree Test suggested four factors. Orthogonal Varimax and oblique rotations by the Maxplane method (Eber, 1966) were performed for the four and five factor solutions. Examination of the rotated factor structures indicated that an oblique four factor solution most nearly conformed to a simple structure. This resulted in meaningful psychological interpretations of the factors.

Factor I most clearly represented a measure of general intelligence.

Factor II was defined as a general index of academic achievement and included the measurement of basic skills related to success in classroom situations.

Factor III was defined by those measures which represented figural creative thinking abilities.

Factor IV was defined by those measures which represented verbal creative thinking abilities.

Results of the factor intercorrelation matrix revealed the following: Factor I showed a moderate correlation with both Factor II (.42) and Factor III (.41) and a low correlation with Factor IV (.07). Factor II showed a moderate correlation with Factor III (.44), but essentially no correlation with Factor IV (.007). A low correlation was obtained between Factors II and IV (.08).

In summary, the verbal figural creative thinking factors emerged as independent dimensions. The verbal creative thinking factor was remarkably independent of the general intelligence and academic achievement factors. Furthermore, while moderate correlations were found among Factors I, II, and III, these factors could be identified as separate and distinct dimensions.

These findings would also seem to lend support to Torrance's (1966) rationale for developing a figural and verbal form of the *Torrence Test of Creative Thinking.*

A FACTOR ANALYSIS OF ACHIEVEMENT, SCHOLASTIC APTITUDE, CRITICAL THINKING, AND LOGICAL REASONING SUBTESTS,

John Follman and *William Miller*, University of South Florida; *David Hernandez*, Florida Tech. University

The purpose of this study was to determine the interrelationships between scholastic aptitude and achievement tests, and critical thinking tests and their underlying factorial structure.

Pearson product moment intercorrelations, principal components factor analysis, and Kaiser Varimax rotation were determined for nine *Metropolitan Achievement Tests (METRO)* subtests, two *School and College Ability Tests Form 3B (SCAT)* subtests, the *Logical Reasoning Test Form A Part II (Logical Reasoning)*, five subtests of the *Watson-Glaser Critical Thinking Appraisal Form YM (WG)*, four *Cornell Critical Thinking Test Form X (Cornell)* subtests, and nine "Interpretation" items from *A Test of Problem Solving Form A ("Interpretation")*. These tests were administered to 812 ninth graders in 12 public schools of three strata: "disadvantaged" non-integrated blacks, "disadvantaged" integrated blacks and whites, and "non-disadvantaged" whites.

All intercorrelations were significant at the .01 level. The aptitude and achievement subtests intercorrelated moderately high, but generally correlated low with all other subtests. *Logical Reasoning* correlated moderately low with all subtests except for .55 with *WG Inference*. Both the *WG* and *Cornell* subtests correlated moderately low with the scholastic aptitude and achievement subtests, intercorrelated moderately high with their respective subtests, generally correlated moderately low with subtests of the other critical thinking test, and low with *"Interpretations"* except for .51 with *Cornell Deduction* and .48 with *Cornell Assumptions.*

Four factors which accounted for 75 percent of the variance were identified and rotated. The largest factor, accounting for 50 percent of the variance, was an achievement factor consisting of loadings above .76 on all aptitude and achievement subtests. The second factor, accounting for 21 percent of the variance, was a *WG* test factor with all *WG* subtests loading above .61. The third factor, accounting for 20 percent of the variance, was a *Cornell* test and "Interpretations" factor with all five loadings above .61. The fourth factor was weak.

It was concluded that the abilities measured by the *SCAT* and *METRO* subtests differ from those measured by the *WG* and *Cornell* critical thinking subtests, and that the abilities measured by the *WG, Cornell,* and *Logical Reasoning* subtests are all dissimilar.

PATTERN OF ACHIEVEMENT AS RELATED TO THE PERCEIVED-SELF,

Mohindra P. Gill, The Ontario Institute for Studies in Education

The present study was designed to ascertain the relationship between pattern of achievement (under-, average, or over-achievement) and the factorially defined dimensions of the perceived-self as measured by the Self-Concept Scale.

The under-, average, and over-achievers were selected from ninth-grade students in Toronto. Scores on the Canadian Academic Aptitude Test were used to develop regression equations for predicting final average marks by school and sex separately. Under- and over-achievers were defined as those students whose final average marks were, respectively lower or higher than their predicted marks by at least one standard error of estimate. Students showing minimum discrepancy were designated as average achievers. Under-, average, and over-achievers were equally represented, with 68 boys and 68 girls in each category. Each of the achievement groups was classified into four ability levels on the basis of Henmon-Nelson IQs.

For the selected students, scores were obtained on each of the eight factors of the Perceived-Self and on all factors combined. Means and standard deviations were then computed for the various groups determined by sex, pattern of achievement, and ability level. The data were further analyzed by applying a three-way analysis of variance.

On all factors except "Originality," the mean scores of the three groups were significantly different; those of over-achievers were invariably highest, followed by those of average and then under-achievers.

Significant differences were found between the mean scores of boys and girls only on the factors pertaining to "Achievement-Related Characterisitcs" (favouring girls) and "Originality" (favouring boys). Students of various ability levels did not show significant differences among their mean scores, except on "Concentrating Ability."

The findings of the study indicate that a student's academic performance is affected by his perceived-self. As such the school could raise the achievement level of under-achievers by providing them with opportunities for success and self expression.

Session 7.12 **Division D**

RESEARCH ON TESTING

Frank B. Womer, University of Michigan

FAKING AND FAKING DETECTION ON THE KUDER OCCUPATIONAL INTEREST SURVEY, *John R. Braun*, University of Bridgeport

The Kuder Occupational Interest Survey (OIS; Kuder, 1966) yields Lamba coefficient scores on 115 ocupational and 33 college-major scales. Earlier Kuder forms have been found fakable (Durnall, 1954; Longstaff, 1948), and the OIS manual gives preliminary data on effectiveness of several means of detecting faking. Research has shown difficulty in establishing faking detection keys that hold up well with new groups (Braun, 1968; Norman, 1963). Also, OIS manual does not indicate how successful *S*s are in modifying scores on pertinent scales other than detection. Therefore, we investigated effects of two faking sets on OIS performance.

*S*s were 32 male and 33 female undergraduates. Each completed the OIS under standard instructions and, one week later, under faking instructions. The instructions for males emphasized giving responses that would help them get into medical school, while those for females involved getting into nursing school. *S*s were told to try to fake so that their faking would not be detected.

It was found that *S*s shifted scores significantly on pertinent scales (e.g., Premedical College Major for men, Nursing College Major scale for women) under faking instructions. *S*s were more successful in altering scores on College Major scales than on Occupational scales. However, *S*s frequently earned high scores on related occupational scales, even though they might have missed on the Physician or Nurse scales, *per se*. Thus not one male had his highest occupational score on Podiatrist under standard instructions, while 72% did while faking. Only 9% of the women had their highest standard administration score on either Dental Assistant, Physical Therapist, X-Ray Technician or HS Science Teacher, while 76% did while faking.

The faking detection keys held up well with our groups; 80% of *S*s under faking instructions earned unacceptable scores on at least one of the detection keys, with only 11% false positives for standard instructions.

THE EFFECTS OF TYPE OF EXAMINATION ANTICIPATED ON STUDENT TEST PREPARATION AND PERFORMANCE, *A. Ralph Hakstian*, Laboratory of Educational Research, University of Colorado

It is commonly contended (Stanley, 1964; Adams, 1965) that students anticipating an essay examination prepare differently (mastering larger units of information) than those expecting an objective test. This study investigated the effects of anticipation of either an objective, combined objective and essay, or essay test on students' preparation procedures and performance on both objective and essay examinations. Two experiments were conducted. In the first, 36 introductory education students were randomly divided into three groups—anticipating an objective, essay, or combined test. All subjects were given an objective and essay midterm examination and a study habits questionnaire. An unannounced retention test was administered two weeks later. ANOVAs of seven questionnaire variables showed no differences between groups in preparation time, organization of material, or techniques employed. With *SAT*, total scores covaried, and ANOVAs of scores on the (1) objective, (2) essay, (3) retention tests, and (4) subtests of items from the knowledge, comprehension, application, and analysis categories (Bloom's *Taxonomy*) again revealed no group differences.

The second experiment, with language arts students, involved a replication of the first, but with much tighter control over the preparation period. Randomly assigned subjects, after being informed of their examination type, read a short linguistics passage, and the investigator unobtrusively recorded actual time spent, amount of underlining, note making, etc. Immediately after preparation, all subjects took both examinations. ANOVAs of four performance criteria (objective and essay tests, knowledge items, comprehension items) and of the study variables again showed no group differences—corroborating the findings of the first experiment.

It was concluded that type of examination anticipated did not affect amount or type of preparation, or actual performance on either examination type (only partially in agreement with findings of Sax and Collet, 1968). The view that the essay examination promotes more desirable studying procedures and higher performance was not supported.

VARIATIONS IN EXAM SOPHISTICATION AMONG COLLEGE FRESHMAN, *Arvo E. Juola*, Michigan State University

Exam sophistication is to be surveyed by the questionnaire method. A total of 36 items is to analyzed. Each of the items were administered to one of four representative 1,000 student samples of freshmen entering Michigan State University as part of a larger survey of cognitive abilities. The National Assessment model is to be used to chart exam taking behaviors. This procedure regards each item as indicative of an acquired skill, and item response frequencies indicate the relative proficiencies of students surveyed. Analyses will be made of all students to secure a composite index of exam sophistication among freshmen. Specified sub-groups will then be analyzed separately to determine systematic variations in these skills. These sub-groups will include classifications based upon types of high schools (size, traditional versus innovative, location as by inner city versus suburb, etc.) and student characteristics (ability, academic success, choice of major, etc.).

The questions on exam sophistication were based upon practices reported in the recent book by the writer, *Examination Skills and Techniques*. The book evolved from an exhaustive review of the research literature on exam-taking procedures, and the practices proposed were documented by research reports, when available, consistent logic, when applicable, or, as a last resort, informed opinion.

Questions which the analysis of data is designed to answer include the following: What is the relative incidence of different elements in exam sophistication among college freshmen from one large midwest institution? Are there differences in exam sophistication among students from different kinds of high schools, or among students who differ in ability, achievement, or personal characteristics? Are diverse elements of exam sophistication positively correlated, or are even exam-wise students influenced by flagrant popular misconceptions such as toward not changing answers, omitting many questions because of fear of the penalty for guessing, pious faith in the laws of chance, and so on.

STUDENT-INSTRUCTOR AGREEMENT IN SCORING AN ESSAY EXAMINATION, *Stanley M. Nealey*, University of Illinois

The aim of the present study was to develop and test a method whereby students could score their own essay examinations.

Subjects were 46 college students enrolled in an advanced psychology course. All subjects took a 50-minute, 11 item, short-answer essay midterm examination. At the next class meeting, each student was given brief written scoring criteria and randomly assigned to rate all 46 answers to one of the 11 items. The 3 or 4 independent student ratings of each item were then pooled. Subsequently the instructor scored the whole examination in the customary manner.

Scores on the total test obtained by pooling student judgments correlated with the instructor's judgments .922. Item means and standard deviations of the instructor's ratings agreed closely with those of the pooled student ratings. Agreement between individual student raters and the instructor was not correlated with the student's total test score or grade point average. In other words, neither course knowledge nor general ability made for better rating.

The multiple ratings of test performance were used to calculate for each student an "answer-ambiguity index" (the mean standard deviation of the ratings). High scores on this index indicated that raters disagreed about what score to assign. Answer ambiguity was negatively correlated with grade point average (-.318), with class in school (-.352), and with total test score (-.485).

The results support the generalization that subjective scoring standards are readily communicable. The method makes possible a vast saving in staff time and very quick feedback to the student. In addition, the multiple ratings provide empirical information about the test and an index of the ambiguity of each student's answers.

Session 8.1 **Division A**

STUDIES OF ORGANIZATIONAL CLIMATE

Jerrold M. Novotney, University of California at Los Angeles, Chairman

ORGANIZATIONAL CLIMATE AND PRINCIPAL-TEACHER COMMUNICATIONS IN CERTAIN OHIO PUBLIC SCHOOLS, *Carl Helwig*, University of Akron

Through Spearman correlation, frequency of principal-teacher communications was correlated with esprit, a subdimension of A. W. Halpin and D. B. Croft's *Organizational Climate Description Questionnaire.* The sample consisted of thirty-seven cooperating Ohio elementary school principals and 310 teachers. The general hypothesis tested was that the frequency of oral and written communications between a principal and his teachers was related to teacher esprit (morale). Since esprit on the *OCDQ,* according to Halpin and Croft, tended also to vary directly with the climate of the school, it was conjectured that the frequency of principal-teacher communications would also reveal some relationships with school organizational climate.

Findings included no significant relationships at the .05 level of acceptance between the frequency of principal-teacher communications and teacher esprit. When principal-teacher communications were separated into principal downward communications to faculty and teacher upward communications to principal, it was found that neither of these subvariables correlated with esprit. Schools with open and closed school climates, the extreme climates on the *OCDQ,* were then singled out into separate groups. The principal behavioral subdimensions on the *OCDQ* in these two groups of thrust, production emphasis, and aloofness, as well as the teacher behavioral subdimensions of hindrance and disengagement, were correlated with principal downward communications to faculty; again no significant relationships were discovered between the latter variable and each of these subdimensions. These subdimensions were assumed to manifest certain communicative styles with the oral and written aspects further assumed to be latent ingredients within these larger behavioral patterns. With similar hypothesizing, teacher upward communications to the principal were correlated in open and closed climate schools with the teacher subdimensions of disengagement and esprit, and again no significant relationships were obtained.

The basic implication seemed to be that the frequency of principal and/or teacher communicative interaction within the elementary school has nothing to do with teacher morale or school organizational climate.

THE FACTOR STRUCTURE OF "THE ORGANIZATIONAL CLIMATE DESCRIPTION QUESTIONNAIRE" FOR TEACHERS IN FIVE URBAN AREAS, *James B. Kenney*, University of Georgia

The structure of perception held by elementary teachers toward the "climate" of their schools was studied in five urban areas. The Organizational Climate Description Questionnaire (OCDQ) was the instrument used in the study. Composition of the sample was as follows: Atlanta, Ga.—102 schools, Baltimore, Md.—67 schools, Chicago, Ill.—78 schools, Houston, Tex. —51 schools, and New York City—80 schools. Only teachers (N=6755) were included in this study.

Data from each city were factor analyzed separately and studied for dimensionality. Factors were compared across groups and identified. The data were then factored as a single group. Dimensionality was examined by graphing root size (Cattell's scree criterion). Six factors appear to be significant on the separate and total factor analyses in both orthogonal and oblique solutions. In order to determine the best possible factor structure, 8, 6, and 4 factor solutions were examined through Varimax rotation (orthogonal) and Maxplane (oblique) rotational procedures. Maxplane was used to determine maximum simple structure. Item loadings and item taxonomies are presented in tabular form and discussed. The scree criterion and inspection of the rotated factors showed that, from both a statistical and psychologically meaningful point of view, the six factor solution was most appropriate. Halpin and Croft identified eight dimensions of the OCDQ in their original study of a heterogenous sample of schools.

The six derived factors were named as follows: (I) Principal-Authority figure, (II) Teacher-qua teacher group perception, (III) Non-classroom School satisfaction, (IV) Principal-teacher communication, (V) Work conditions, and (VI) Availability of school service and materials. Factors III, IV, and V closely resemble three factors obtained by Rentz and Kenney in development of the Teacher Opinion Rating Scale (TORS). Item communalities for the six factor solution were, in virtually all instances, higher than those presented in the Halpin and Croft study.

The notion is advanced that the OCDQ, at the sub-test level, is measuring a different kind of teacher perception of the school in a completely urban setting than was the case in the original Halpin and Croft factor analysis. There appears to be some evidence that, beyond what the items literally say, perceptions apparently unique to the urban school are being projected into their interpretation by teachers. Whatever it is, this perceptual variance results in realignment of the factor structure and appears to be associated with the situational context within which the teachers work. It is not a product of the organization; nor is it idiosyncratic to the individual.

LEADER BEHAVIOR CHARACTERISTICS AND ORGANIZATIONAL CLIMATE, *Thomas W. Wiggins*, University of Oklahoma

The purposes of this study were to investigate leader behavior characteristics of elementary school principals and to examine these characteristics as they relate to the organizational climate. The initial interest was the conceptualization of leader behavior and organizational climate, which was accomplished by means of the theoretical construct of social systems theory. Specifically, the social systems model provided the justification for (1) viewing the leader's behavior characteristically as his need-dispositions (or values) and orientations, and (2) conceptualizing organizational climate as the interaction between the task-achievement and the needs-satisfaction dimensions within the organization.

The general hypothesis of the study was that there exists a significant relationship between leader behavior characteristics of elementary principals and the organizational climates of the schools within which they serve. Specific hypotheses dealt with (1) the effect of the replacement of a principal upon the organizational climate, and (2) the effect of the length of a principal's incumbency in a school upon both his leader behavior and his perceptions of the organizational climate.

Data on the organizational climate were collected by means of the *Organizational Climate Description Questionnaire* (*OCDQ*). Data on leader behavior characteristics were collected by means of the *Fundamental Interpersonal Relationship Orientation-Behavior* (*FIRO-B*) questionnaire, *Orientation Inventory* (*ORI*), and the *Survey of Interpersonal Values* (*SIV*). The unit of analysis was the school. The sample consisted of thirty-five randomly selected schools in one large urban school district in Southern California. The statistics used were canonical correlation analysis and trend analysis utilizing the "t"-test of significance.

Generally, leader behavior and organizational climate were not shown to be significantly related. However, a significant relationship was revealed between the interpersonal orientation (*FIRO-B*) of the principal and the organizational climate (*OCDQ*). The stability of the organizational climate was suggested by both the overwhelming support for the hypothesis dealing with principal replacement, and by the investigation of five specific plausible explanations. The replacement of the principal was shown to have no significant effect upon the existing organizational climate. The length of the principal's incumbency was shown to be related to the congruence of his leader behavior and organizational climate, but not related to the congruence of the teacher-principal perceptions of their organizational climate.

Implications for further research include a more precise consideration of the construct validity of the instruments used, reanalysis of the data to compare principal and teacher perceptions of the domains of interest exclusive of one another, an investigation into the possible etiology of climate stability and leader behavior uniformity, and further development of canonical correlation analysis with particular emphasis upon normalization of the tests and significance for the coefficients.

PERSONAL VARIABLES OF TEACHERS AS RELATED TO THEIR PERCEPTION OF DIMENSIONS OF ORGANIZATIONAL CLIMATE, *Eldon J. Null*, Purdue University

Research has provided evidence that an individual's perception of the behavior of other people is related to his own pattern of personal variables. Further, this relationship is often associated with identifiable perceptual errors. The purpose of this study was to determine if relationships exist between personal variables of teachers and their evaluation of the behavior of their principal and fellow teachers in responses to items of the *Organizational Climate Description Questionnaire.*

A stratified random sample of thirty-one Minneapolis and St. Paul schools and fifty suburban schools was drawn from a population of 154 elementary schools which met certain predetermined criteria. Each of the 1,691 teachers in this sample responded to the OCDQ and to either the *Minnesota Teacher Attitude Inventory* or the *Sixteen Personality Factor Questionnaire* as determined by random chance.

Eight coefficients of correlation, calculated to relate MTAI scores to each of the OCDQ subtest scores, were quite low. However, each gave evidence that teachers with a "good" attitude toward children tended to view all eight dimensions of climate in a manner indicative of an open climate, while teachers with a "poor" attitude toward children tended to view all eight dimensions in a manner indicative of a closed climate.

Eight multiple regression equations in which the 16 PF Questionnaire scores served as predictors of the OCDQ subtest scores, yielded twenty-two significant (.05 level) beta coefficients of the 128 tested. While a number of the significant coefficients were quite low, in no instance did a score in a particular personality factor predict one subtest score in a manner indicative of an open climate and a second subtest score in a manner indicative of a closed climate.

The findings of the study suggest that personal variables of respondents should be taken into consideration in an interpretation of OCDQ scores.

ORGANIZATIONAL CLIMATE AND THE ADOPTION OF EDUCATIONAL INNOVATIONS, *Homer M. Johnson*, Utah State University; *R. Laverne Marcum*, Utah State University

The purposes of this study were to determine (1) if there were differences between each of four variables (expenditures, age of staff, years in the school, and number of professional staff) for the most innovative schools and least innovative schools participating in the study, (2) if differences existed between the teacher and administrator perception of school climate for the most innovative schools and for the least innovative schools, and (3) if differences existed between each of the four variables noted above for the open climate and closed climate schools.

Fifteen of the most innovative schools and fifteen of the least innovative schools were identified by means of the Educational Innovation Checklist. The degree of openness and closedness was determined by administering Halpin's Organizational Climate Descriptions Questionnaire. At the same time fiscal biographical data were gathered.

An analysis of variance and F ratio were used to determine if significant differences existed between the means of the four variables and two climate categories for the most innovative and least innovative schools. The same analysis was used to test for differences in teacher and administrator perception of climate in the two innovative categories.

The following results were obtained:

1. Significant differences (.05 level) were found between the most innovative and the least innovative schools in the following variables: climate, expenditure, age of professional staff, years of service of the professional staff, and the number of professional staff.
2. Significant differences (.05 level) were found between the teachers and administrators in their perception of the school climate in the most innovative schools.
3. No significant differences were found between the teachers and administrators in their perception of the climate of the least innovative schools.
4. Significant differences (.05 level) were found between the open climate schools and closed climate schools for the variables of innovation, age of professional staff, and number of years in school for the professional staff.
5. No significant differences were found between the open climate schools and the closed climate schools for the variable of expenditure.

The following conclusions were drawn:

1. Innovative schools have more open climates, higher expenditures, younger staffs, greater staff turnover, large number of staff.
2. In the most innovative schools administrators viewed the climate as more open than did teachers, while in the least innovative schools both groups viewed the schools as closed.

3. Open schools had younger staffs, and the professional staff remained in the school fewer number of years and were more innovative.
4. The higher level of expenditure did not influence the school climate toward openness.

SOCIAL STRUCTURE IN RELATIVELY CLOSED AND OPEN SCHOOLS, *Robert J. Coughlan*, Northwestern University

This study was concerned with examining the ways in which the work values of teachers affect group development within two different types of school organizational systems.

Utilizing Homans' interactionist schema, Barnes suggested the terms "closed" and "open" to describe the external system of a work group in formal organization. An objective procedure was developed for investigating prescribed structural arrangements in the external system of a teacher work group. The teacher groups in 11 suburban high schools matched on technological, economic, and social factors were ranked on 38 external system variables. The teacher groups with the two most closed and the most open external systems became the research sample.

A syndrome analysis of teacher work value data, obtained from a 45-item audit and content analysis on teacher achievements and aspirations, provided the bases for identifying two types of Professionals, two types of Organizationals, one Social type, and a group of Marginals among the 192 teachers in the four sample schools.

Three categories of interpersonal choice—consultation, attraction and respect—were selected as indicators of the network of relations in the teacher work group of each school. Choice data were obtained through the use of a sociometric questionnaire. The influence of work values on interpersonal choices was traced through the chi-square procedure in order to determine if value-orientations have a differentiating, segregating, or no effect on the group's social structure.

The findings from both the relatively closed and open schools were consistent with a revision in understanding of the conceptual scheme underlying the research. In the relatively closed schools, high status and prestige were bestowed on those Professionals whose values and behavior were most supportive of internal system development. In the relatively open schools, informal group leadership gravitated to those Organizationals whose values and behavior were most supportive of fused external-internal system development.

Session 8.7 **Division C**

ACQUISITION OF READING SKILLS NO. 2

William Eller, State University of New York at Buffalo, Chairman

CHILDREN'S WORD ASSOCIATIONS TO INDIVIDUAL LETTERS, *Mary Lou Ellison* and *Wayne Otto*, Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, Madison

Children's single word responses to visually presented upper and lower case forms of the letters of the alphabet were examined. There were two basic purposes: first, to compare the children's responses to adults' responses reported in earlier studies; and second, to answer specific questions regarding children's responses. The predictions were that (1) girls would give more consensual responses than boys, (2) responses to upper and lower case letter forms would not differ, (3) the most frequent responses would come from high frequency words on the Lorge-Thorndike list, and (4) the ranking of letters by frequency of consensual first responses would be negatively related to the ranking of letters by the total array of words beginning with each letter.

Thirty-five boys and 35 girls from the sixth grades of each of two elementary schools participated in the study, making a total of 140 subjects. Each subject was given a 26-page booklet with one letter printed at the top of each page and told to "write the first word you think of that begins with the letter printed at the top of the page." The children from one school responded to lower case letter forms and the children from the other school responded to upper case letter forms.

The children's responses were, in general, similar to the previously reported responses of adults. The responses of boys and girls and the responses to upper and lower case letter forms were generally similar. Frequently given responses tended to come from among high frequency words on the Lorge-Thorndike list, and there was evidence that the tendency to respond consensually to letter stimuli is strengthened when the array of available responses is limited. The results are discussed in relation to previous research and in terms of implications for further work.

THE NEW CASTLE STORY, *Robert B. Hayes*, Department of Public Instruction, Harrisburg, Pennsylvania

This was a three-year study in New Castle, Pennsylvania, of the following beginning reading approaches: (1) initial teaching alphabet (i/t/a), (2) Lippincott, Scott, Foresman, and (3) Phonics and Word Power. When i/t/a pupils made transition to traditional orthograph, the Merrill Treasury of Literature Series was employed. The population was randomly selected by attendance areas and assigned to classrooms and treatment groups prior to first grade. The study began with 415 first grade pupils and twenty classrooms with five classrooms per reading method.

Teachers received frequent in-service education and supervision to insure that they appropriately used the reading approach to which they were assigned. Each year the results were evaluated by the Stanford Achievement Test, the San Diego County Inventory of Reading Attitude, and the number of books read independently. In addition, randomly selected subsamples were given the Gilmore Oral Reading Test, the Gates Word Pronunciation Test, and the Fry Phonetically Regular Words Oral Reading Test.

The results indicate that intensive phonic methods and materials which introduce many different words can make a significant difference in increased pupil achievement. Retention rates for the low IQ third suggest that a slower instructional pace (than was employed in this study) is advisable for intensive phonics, heavy vocabulary programs. Basal "whole word" readers probably should increase their vocabulary and phonic components.

CONSENSUALITY OF CHILDREN'S WORD ASSOCIATIONS AND FORMULATION OF MAIN IDEAS IN READING, *Karl Koenke*, University of Illinois, Urbana; *Wayne Otto* and *Judith Schein*, University of Wisconsin, Madison

Findings of word association studies with various groups of children and adults suggest that successful reading may be, in part, a function of the reader's consensuality of response or cue elaboration. Since consensuality seems to be a factor leading to successful reading comprehension, and the ability to state a main idea is a comprehension skill, a relationship between children's consensuality scores and their ability to formulate and state a main idea in reading is expected.

Fifty boys and fifty girls from second grade and identical numbers from fifth grade had two tasks: (1) to read three paragraphs of first grade readability and state the implicit main ideas of each, and (2) to associate freely to selected words from the Kent-Rosanoff list. The main idea statements were rated on a six-point quality scale and on a thirteen-point descriptive-structure scale. The scales were the products of two years of continuous study by a committee of researchers of children's ability to conceptualize the main idea. The word associations were scored for frequency of appearance among the responses of the whole sample.

Results show that, as previously reported, there is a significant relationship between reading ability and consensuality of word association/cue elaboration. On the other hand, there seems to be little relationship between consensuality and formulating and stating the implicit main idea of a paragraph. The paragraphs were written to be read by first graders and therefore contained a simplified vocabulary, which cannot be said of standardized reading tests. Thus, the results are discussed in terms of the possible relationship of consensuality of word association to vocabulary level of material and hence vocabulary tests.

LONGITUDINAL EFFECTS OF I.T.A. ON PUPILS' READING ACHIEVEMENT THROUGH GRADE THREE, *Wai-Ching Ho, Charles Eiszler* and *Vickie Stroh*, Educational Research Council of America

Two major purposes of this study are: (1) investigation of longitudinal effects of i.t.a. (initial teaching alphabet) on reading achievement of pupils of various ability levels; (2) feasibility of starting i.t.a. instruction in kindergarten.

The subjects included approximately 600 pupils from 9 school districts. About one third of the pupils started T.O. (Traditional Orthography) in grade 1, another third started i.t.a. in grade 1, and the remaining third started i.t.a. in kindergarten. Reading achievement was measured by the reading tests in the Standard Achievement Battery: Word Reading, Paragraph Meaning, Vocabulary, Spelling, and Word Study Skills for grade 1; Word Meaning, Paragraph Meaning, Spelling, Word Study Skills, and language for grades 2 and 3. In the fifth month of grade 1, pupils took the tests in their own medium of instruction. At the end of grades 1, 2, and 3, all pupils were tested in T.O.

The three groups were compared on each test by ability level for all four testings. The ability levels were determined by the Lorge-Thorndike IQ: Low (ranging from 64 to 103), Middle (104 to 114), and High (115 to 139). For each comparison, a two-way hierarchical analysis of variance (the nested factor design) was employed to isolate the unique effects associated with each classroom.

Four major findings emerged from the analyses which were generally true for all ability levels: (1) i.t.a. pupils generally demonstrated an early superiority over T.O. pupils in Word Reading, Spelling, and Word Study Skills; (2) no significant difference was found between the i.t.a. pupils and T.O. pupils on most Stanford tests beyond grade 1; (3) pupils who started i.t.a. in kindergarten demonstrated consistently better achievement than the other two groups over all three years; (4) spelling did not seem to cause particular difficulty for i.t.a. pupils after they transferred to T.O.

Session 8.8 **Division B**

CONTRIBUTED PAPERS NO. 1

Chairman: To be announced

CURRICULUM AND ORGANIZATIONAL PATTERNS OF MIDDLE SCHOOLS, *William M. Alexander*, University of Florida

This research project was undertaken to provide bench-mark data regarding middle schools. A list of 1101 schools, each of which met the survey definition of "a school which combines into one organizational facility certain school years (usually grades 5-8 or 6-8) which have in the past usually been separated in elementary and secondary schools under such plans as the 6-3-3, 6-2-4, and 6-6," was compiled from state departments of education and other sources. A survey instrument was used to secure detailed data from a 10 percent random sample stratified by USOE regions, and schools were visited for supplementary information.

Of the sample, only 10.4 percent had been established before 1960. It is clear that there is a current movement toward reorganization of the school ladder, with grade 6-8 (60 percent of the sample) and grade 5-8 (27.3 percent) schools predominating in the reorganization. Other data from the survey, however, indicate that only the grade patterns are different in most of the schools, with the program of studies and organizational patterns closely resembling those of the predecessor organizations, especially the grade 7-9 junior high school. Some potential for newer curriculum and organizational patterns seemingly appropriate to middle school learners was identified in a few new schools. Generally, more careful planning and evaluation of the new organizations seem essential.

THE RELATIONSHIP OF FORM OF SCHOOL ORGANIZATION TO PUPIL BEHAVIOR, *Francis X. Vogel*, Northeastern Illinois State College and *Norman D. Bowers*, Northwestern University

This study explored the relationship of the form of school organization to pupil attitudes, achievement, conceptual maturity, and classroom behavior. Ten teachers in each of three groups in one school system were selected by a random process, and the 707 pupils enrolled in their classes constituted the sample. The experimental group was placed in a nongraded form of organization. The control groups represented the traditional (graded) form of organization. The pupils in all groups were subdivided into three age groups: normal age, underage, and overage.

Data were collected from the pupils in the fall, winter, and spring of one school year. The specific measures included the *Describe-Your-School, Stanford Achievement Tests,* the *Draw-A-Person,* the *Russell Sage Social Relations Test,* and two scales derived from the *Observation Schedule and Record* (*2e*). The scores obtained in the fall on the *Describe-Your School, Stanford Achievement Test,* and the *Draw-A-Person* were used as covariates; data collected subsequently were the dependent variables.

Multivariate analyses of covariance were completed to determine the interaction of the age groups and school groups, the differences among the age groups, and the differences among the school groups. Significant differences ($p < .01$) were found for the interaction among the age groups and among the school groups.

Univariate analyses performed to indicate directionality of the differences among the age groups revealed that the underage pupils had the highest scores ($p < .01$) on the measures of achievement, group planning, and conceptual maturity. For the scale of group operations-contributing, the overage pupils were the highest ($p < .05$). For the observation-contributing scale, the normal age pupils scored the highest ($p < .01$). For all other measures, no differences were found that were statistically significant.

The univariate analyses utilizing measures on the school groups showed higher scores for the experimental group ($p < .01$) on measures of conceptual maturity, group planning, and the observations-non-contributing scale. Control group one had higher scores on measures of achievement ($p < .01$), attitudes ($p < .01$), and the observations-contributing scale ($p < .05$). Control group two had the highest scores on the operations-contributing scale ($p < .01$). There were no significant differences among the groups on the measure of group operations-non-contributing scale.

Interpretation of data and an analysis of usefulness of this design for this type of problem are discussed.

A STUDY OF A FIXED SEQUENCE OF SKILL AND CONCEPT ACQUISITION REQUISITE TO PERFORMANCE OF A COMMON SCHOOL TASK: MAP DRAWING, *David H. Feldman,* Stanford University

The purpose of the present study was to gather empirical evidence bearing on the hypothesis that a fixed sequence of concept and skill acquisition is requisite to representation of space in the drawing of a proper geographic map.

Using Piagetian theory and observations as a guide, modifications in the construction and administration of a map-reading test, first developed by Salomon (1968), were made so that the test items were arranged in a hypothesized fixed sequence. A map-drawing test was also administered.

On the basis of data gathered for 46 fourth, fifth, and sixth grade *Ss* it was found that: a) there was a correspondence between the median reasoning level *Ss* reflected on the map reading test and their map drawing category; b) map reading scores and map drawing categories were correlated .73 ($p < .001$), supporting the hypothesis that the two tests demanded the same concepts and skills; c) a positive but nonsignificant correlation between map reading score and IQ supported the hypothesis that spatial reasoning and not reading was being measured; d) Salomon's three kinds of questions were found in the present study, replicating earlier results;

and e) a Subjects x Items pass/fail analysis revealed that nearly 75% of the items in the map reading test were distributed as predicted by the fixed-sequence hypothesis.

Design and interpretation problems in the study were discussed, future research outlined, and educational implications, especially *vis a vis* educational programs for the "disadvantaged," were drawn.

IMAGES OF AFRICA: WHAT AMERICAN SECONDARY SCHOOL STUDENTS KNOW ABOUT AFRICA SOUTH OF THE SAHARA, *Dr. E. Perry Hicks,* State University of New York at Buffalo and *Dr. Barry K. Beyer,* Carnegie-Mellon University

Preparation of an effective curriculum requires an awareness of the knowledge and impressions already held about the topic of study by prospective students. Accordingly, the staff of Project Africa felt that the initial step in its proposed preparation of instructional materials should be a determination of what American secondary school students know and believe about Africa south of the Sahara.

A total of 3259 average-ability seventh and twelfth graders in 28 school systems across the United States were tested using two different instruments. The first consisted of an outline map of the world and a deck of stimulus cards. Students matched each of the ninety stimulus cards with the area on the map which it seemed to describe best. The stimulus cards contained words such as *friend, cowardly, religion, villages, superstition,* etc. The second instrument tested students' knowledge about six aspects of Africa south of the Sahara: Physical Geography, History before the European Penetration, History of the Europeans in Africa, Indigenous Society, Economic Development, and Current Affairs.

The results indicate that neither seventh grade nor twelfth grade students are very knowledgeable about Africa south of the Sahara, and that much of what students believe to be true is based on certain widely-held misconceptions. Their concept of this region is one of primitive backwardness, steaming jungles, strange peoples, danger, and savage violence.

Curriculum planners need to be cognizant of this image of Africa south of the Sahara. Class presentations, whether these be through lectures, films, or textbooks, should present a balanced view of the African scene. Perhaps there should even be an attempt to destroy misconceptions by presenting material which contradicts, rather than reinforces, what students think or believe about this region.

DEVELOPING A COGNITIVE CATEGORY SYSTEM FOR ANALYZING CLASSROOM DISCUSSION ON SOCIAL ISSUES, *Byron G. Massialas, Nancy Freitag* and *Jo A Sweeney,* University of Michigan

The purpose of this study was to design and test a cognitively oriented category system for use in analyzing classroom discourse centered around controversial social issues. The instrument was developed as part of a larger project, *Inquiry into Social Issues,* which is investigating the teaching of controversial issues in secondary schools in the state of Michigan.

Fourteen social studies classes and one English class were selected from a state-wide probability sample of secondary school teachers. Two or three tape recordings of discussions of social issues in each of these classrooms were transcribed and analyzed. A theoretical model was developed to form the basis of the category system; then a sample of the obtained tapes and transcripts was used to modify and further refine the system. Finally, another set of tapes and transcripts was analyzed using the modified category system. The transcripts were analyzed by trained coders working in pairs. All verbal statements were classified by each coder and a series of numbers generated, intercoder reliability coefficients for each transcript were obtained, and the coded series of numbers produced for each transcript were subject to computer matrix analysis.

The resulting category system focuses on the cognitive aspects of both teacher and student interaction in classrooms where social issues are discussed. The unit of measurement (referred to as a cognitive event) focuses on cognitive rather than affective or procedural operations. There are three general sets of categories: categories indicating a speaker's request that a cognitive operation be performed, categories indicating that a speaker is performing a cognitive operation, and categories which take account of non-cognitive events occurring in the classroom. The cognitive categories include introduction, elaboration, summarizing, definition, position statements, and grounding statements. The non-cognitive categories include directions and classroom maintenance, acceptance or encouragement, and negative feedback.

Through analysis of the quantitative patterns of interaction, a clearer understanding takes place of the value judgments and intellectual operations involved in classroom examination of social issues. The analysis of these patterns of interaction may aid the curriculum theorists in understanding the dynamics of cognitive interaction in the classroom.

A MODEL FOR TEACHING COMPOSITION, *Lester S. Golub*, University of Wisconsin

The theory among some teachers of English that writing cannot be taught makes for unimaginative and unorganized teaching of composition. A model for teaching composition has been shown to have some research and teaching value.

The traditional composition model in the United States has been the write-correct-rewrite model which lacks in teaching flexibility and creativity for both students and teachers. It emphasizes written composition without permitting the student to exercise his oral discourse ability, the student's inner voice needed for expression. This paper attempts to explain and show the results of teaching composition from a new model for teaching composition.

Once this model is established and understood, teachers can lead their students through the various phases of the model and can order their instruction in a more systematic way than is possible with the traditional composition model. Also, researchers can test the effects of certain kinds of instruction, activities, and stimuli at various phases of the model. The research described in this paper tested the effects of the uses of oral discourse in the composition teaching model.

The control group of 56 and an experimental group of 56 homogeneously grouped ninth-grade students were matched for IQ, socioeconomic background, and age. The research was conducted in thirteen days. The control group went from *stimulus* to *problem* to *written discussion* on the traditional model, the experimental group went from *stimulus* to *problem* to *controlled oral discussion* to *written discussion.* Three teachers evaluated, on a seven-point scale, unidentified papers of both pretest and posttest of control and experimental groups. With the use of the chi square statistic it is possible to show that the use of controlled oral discourse produced more grade increases and fewer decreases in the experimental group than the control group at the .001 level of significance.

The use of the model for teaching composition has a pedagogical value in offering alternatives to the traditional model for teaching composition. Experimentally, the model offers a systematic approach to research which can test the effects of teaching-learning activities suggested in the model.

SOME INNOVATIVE APPROACHES TO THE TEACHING AND EVALUATION OF WRITTEN COMPOSITION, *Sarah D. Hervey*, University of Arizona and *Clinton S. Burhans, Jr.*, Michigan State University

A series of USOE-sponsored Project English studies at Michigan State University has explored various philosophical approaches to the teaching of composition. Earlier phases of Project English identified and explored the use of two approaches: (1) *Pre-writing,* centering particularly on the formative stages of the writing process, and (2) a more *Comprehensive* approach, which added to the concept of pre-writing some new tactics for the writing process itself.

The most recent and most comprehensive Project English study was an experiment designed to compare in a formal way these two innovative approaches with a traditional approach. The experiment also included a comparison of two course formats: (1) the usual three hours' weekly class of thirty students, and (2) a tutorial experience including one lecture and one tutorial session weekly. A 2x3 groups-within-treatments design with random assignment of students to groups (classes) and rotation of teachers across treatments was used.

An adaptation of Diederich's E.T.S. scale was used by nine readers whose ratings provided the basic data for the analysis. The readers, master high school English teachers, were trained with the scale before evaluating the papers; interjudge reliabilities ranged from .52 to .85.

Analysis of the treatment effects show that *Pre-writing* produced better wording ($p = .056$) and flavor ($p = .004$), and *Comprehensive* produced better expression of ideas ($p = .044$) and organization ($p = .035$); in the mechanics of writing none of the three approaches was superior. Neither course format proved to be clearly superior.

It was concluded that fresh and vital, more personal and creative approaches can produce clearly better writing, and that the techniques traditionally used in high school and college could be drastically revised, if not abandoned altogether. The two approaches tested in this study show promise of providing a basis for substantial and widespread revision of composition instruction.

OBJECTIVES FOR COMPENSATORY PROGRAMS, *Ralph H. Locklin, Barbara Harrison* and *Forest I. Harrison,* Claremont Graduate School

A USOE-sponsored international survey was conducted to determine the objectives for compensatory programs for culturally disadvantaged adolescents.

Papers disclosing objectives were requested from 322 educators, 173 of whom responded. From 432 sources, including the responses, a list of 1,814 objectives was compiled, analyzed, and classified into three categories: student objectives (57%), program activities (37%), and uninterpretable (6%). The student objectives (1,041) were classified in a behavior-by-content 2 x 4 array. These objectives were sorted by cognitive (54%) and affective (46%) behaviors based on the *Taxonomies.* They were further sorted by content: subject area (38%), general academic development (29%), general social development (27%), and vocational development (6%). Of the objectives within general academic development and the objectives within vocational development, about 45% involved cognitive behaviors. Within subject area, however, over 80% of the objectives involving cognitive behaviors, and within general social development, only 24% were cognitive. Finally, 73% of the subject area cognitive development objectives dealt with language development.

Educators concerned with compensatory educational programs are formulating objectives selectively with respect to both content and behavior. The results of this study indicate that educators are placing priorities on objectives as follows: subject area cognitive development, affective general social development, affective general academic development, cognitive general academic development, subject area affective development, cognitive general social development, and, lastly, cognitive and affective vocational development.

ADOLESCENT SOCIETY REVISITED: THE SOCIAL-ACADEMIC-PERSONAL EXPERIENCES OF PUPILS, *James H. McElhinney* and *Richard C. Kunkel,* Ball State University and *Donald G. Ferguson,* New Mexico State University

Quality curriculum assessment needs to be unified by a theory. The authors theorized that to plan curriculum change intelligently it is valuable to know the kind and the significance of the social-academic-personal experiences pupils now have because they attend school. The purpose of this study was to collect parallel data with multiple instruments from pupils, teachers, administrators, and instructional service personnel.

Quality assessment should be based on the research of others and structured by an appropriate design. Coleman used questionnaires and solicited responses of pupils. Ryans collected data from teachers, administrators, and others. Using their work, supplemented by other research, focused personal interviews and parallel questionnaires were used with 3000 adolescents, 900 teachers, and 100 administrators and service persons in thirteen school systems in East Central Indiana. Teachers and other professional personnel reported on what happened to pupils and made judgments on how and why these experiences were provided. Pupils reported on what was happening to them and how they perceived it.

Data were collected on pupil-teacher relationships, learning and instructional practices, pupil-pupil relationships, social and personal activities. Information accumulated on a number of topics including pupil reading patterns, television viewing habits, amount of autonomy of pupils and teachers, pupil-teacher agreement and disagreement, applicability outside of school, vocational and educational planning, factors in self-concept and in other areas.

The theory design makes two basic assumptions. One assumption is that education is what pupils perceive is happening to them because they attend school. The second assumption is that education is what teachers (school personnel) do that influence pupils. When data are combined from all sources and coupled with an analysis of the world in which these pupils live, the information gathered in the assessment provides major implications for the revision of curriculum.

CURRICULAR INNOVATION AND THE SUBJECT-MATTER PREPARATION OF COLLEGE BOARD TEST CANDIDATES, *Elizabeth W. Haven*, Educational Testing Service

The widespread movement toward curricular innovation which began in the late 1950's created the need for up-to-date information on student learning experiences to aid in developing achievement tests for college-bound secondary students. Accordingly, testing specialists from Educational Testing Service, working with committees of examiners in each subject, formulated ten questionnaires covering the main academic subjects taught in secondary school.

Questionnaires were mailed in the spring of 1966 to 38,114 students from 7,555 high schools throughout the country who had taken at least one College Board achievement test during the 1965-66 academic year. Samples of equal size were drawn for all tests within an administration. Because no student was to be included in more than one sample within an administration, the selection method was not strictly random. The few students who happened to be drawn for two samples were excluded from the sample for the more popular test.

The discussion of results here is limited to three subjects: mathematics, French, and biology. There is unmistakable evidence that the innovations recommended by the Commission on Mathematics and other curricular reform groups were integrated into the mathematics programs of these students, some of the concepts being taught as early as seventh grade. Students in the French sample reported that their introduction to the language, even in the early 60's, was principally through listening and speaking and that the language laboratory was commonly used. However, they also reported substantial emphasis on preparing written translations and using English to explain vocabulary or grammar. The Biological Sciences Curriculum Study (BSCS) movement, although very recent, exerted tremendous influence on what these students reported studying in biology. While the differences in emphasis among these courses of study were consistent with textbook content, the results showed that a great deal of nontextual material was also included in all biology curriculums.

SURVEY OF INNOVATIONS IN JUNIOR COLLEGE CURRICULUM AND INSTRUCTION, *B. Lamar Johnson*, University of California at Los Angeles

The purposes of the survey were to identify promising innovations in the junior college curriculum and instruction, to identify conditions which encourage such innovation, and to project trends and recommended practices designed to change and improve the instructional program.

Findings are based on visits to seventy-seven junior colleges in twenty-two states, and on conferences with representatives of an additional eighty-two colleges in twelve of these states. Findings from these visits and conferences were augmented by written reports from more than one hundred colleges in twenty-seven states and the District of Columbia.

Course content and organization were analyzed, as were also examinations and instructional materials. Classes were observed, and interviews were held with administrators, teachers, and students. Particular attention was given to the examination and analysis of plans which utilize the definition of specific objectives as the basis for curriculum development and the improvement of instruction.

The survey identified considerably more new developments than the author found in a similar, though more limited, survey in 1963.

Plans used by junior colleges to encourage innovation include (1) establishing positions or committees to provide leadership for innovation and experimentation, (2) providing innovation centers or libraries, workshops, and conferences, and (3) faculty surveys of educational innovations, including visits to colleges where they may be observed.

Practices which were reported include audio-tutorial teaching, programmed instruction, honors programs, independent study, systems approaches to instruction, student operated experimental colleges, and the use of varied technological aids to learning.

There is a paucity of evidence regarding the outcomes of the various plans which were reported.

A STUDY OF AN INTERDISCIPLINARY ENQUIRY CURRICULUM IN FOUR ENGLISH SECONDARY MODERN SCHOOLS, *James B. Macdonald*, University of Wisconsin-Milwaukee and *Sam Mauger*, Goldsmith College, University of London

The purpose of the study was to analyze the problems and procedures of an Interdisciplinary Enquiry Curriculum operating in four schools in the Greater London area.

Two investigators were invited by the schools to spend three days each in four separate secondary modern schools in the London area, on a consultant-participant-observer basis, working with teams of teachers involved in Interdisciplinary Enquiry programs.

Problems and procedures were identified by observation and interview (both individual and group) with teachers and pupils, and positive and negative reactions to curriculum experiences were gathered from both students and staff via the same methods.

Results revealed that youngsters in these programs, as a whole, were very definite and positive about the quality of their experiences. Teachers, on the other hand, were more ambivalent. Major problems arose in the definition of the teachers' role when students were in Enquiry, and in the definition of the Enquiry process itself. An additional problem was the ability of both teachers and pupils to avail themselves of standards for assessing progress in the program.

Analysis of this program suggests to the investigators that the Interdisciplinary Enquiry approach to curriculum could well have great potential for some of our urban inner city schools and as a basic Middle School Curriculum in America.

SOME REACTIONS OF KINDERGARTEN CHILDREN TO INTELLECTUAL CHALLENGES, *Dr. Doris Fromberg,* Hofstra University

Major purposes for studying kindergarten children's reactions to intellectual challenge included finding out (1) what verbal and observable behavior children manifested, (2) if a relation existed between children's exposure to intellectual challenge and their estimate of challenge, and (3) if they manifested any unique perceptions of the school situation.

The treatment group, consisting of twenty-six five-year-olds in a private New York City kindergarten, was exposed to inductive learning activities connected with a semester's science curriculum. Pre-tests and post-tests, which were administered to the non-treatment group and the treatment group, consisted of an "Estimate of Challenge Task" interview, an interview concerning their perception of the school situation, and the Goodenough "Draw-A-Man" Intelligence Test.

Daily episodic recordings made of the structured and unstructured intervening activities of the treatment group were analyzed by utilizing researcher-developed categorizations of children's verbal and observable behavior.

The treatment group manifested no measurable negative reactions to the intervention program. Increased self-confidence was indicated by their performance on the "Estimate of Challenge Task." They uniquely perceived the school situation, especially the teacher as a resource for ideas as well as physical skills and the distinct difference between kindergarten and nursery school. During intervening experiences, treatment group children appeared receptive to several planned inductive teaching techniques. The relation of physical and verbal communication among children and between children and teacher during structured and unstructured situations was described.

It is implied that teacher-structured, inductively organized intellectual challenge could be expected to help bolster young children's self-confidence. Provision for physical environment with concrete materials in a potentially social environment and such involvement as a potential means for evaluating education programs seem indicated. Early childhood curricular planning based upon conceptual and methodological consistency with disciplines seems supported.

SCIENCE AND MATHEMATICS INSTRUCTION IN KINDERGARTEN AND FIRST GRADE: OUTCOMES IN LOGICAL THINKING IN SECOND GRADE, *Millie Almy,* Teachers College, Columbia University and *Lilly Dimitrovsky,* Child Development Center, New York City

Do children who receive systematic instruction in the basic concepts of mathematics and science when they are in the kindergarten think more effectively when they reach second grade than do children who did not have such early instruction?

One hundred thirty-six children who had no prescribed program of instruction in mathematics and science in either kindergarten or first grade were compared with 294 children who received instruction in either AAAS or SCIS science and GCMP mathematics, or in mathematics alone from kindergarten on. These groups were also compared with 334 children for whom prescribed instruction in science and mathematics, or mathematics alone, began only at first grade.

The measure of logical thinking was based on individual interviews derived from Piaget. At the second grade level, seven kinds of tasks involving conservation, serial ordering, and classification were included. Children were scored operational (giving clear and consistent evidence of logical thinking) or not operational in each of the seven tasks. Performances in the tasks were not highly correlated.

Overall the results highlight the extent of individual variation and the difficulty of using curricular modification as a means of changing basic ways of thinking. However, differences in the performances in specific tasks do reflect to some extent the emphases in the instructional programs.

AN EXAMINATION OF THE INTERDEPENDENCE AMONG VARIABLES RELATED TO SPECIFIED PREDISPOSITIONS AS MANIFESTED THROUGH POPULAR READING MOTIFS IN THE GRAMMAR GRADES, *Joseph Carlton Johnson II*, Duke University and *Milton D. Jacobson*, University of Virginia

The purpose of this investigation was to examine the reader in the upper grammar grades with respect to the manner in which his psychological set, as determined by Likert-type attitudinal appraisal devices, toward specified motif content is related to his comprehension of reading selections reflecting such content. Specifically, the procedures for this study were such as to compare the literal and interpretative comprehension abilities of intermediate grade children on certain thematically based reading selections with their attitudes toward that material.

The investigation was designed to control the predictor variables of race, sex, intelligence level, school, grade, attitude, literal reading comprehension, interpretative reading comprehension, socioeconomic status, and chronological age and the criterion variables of general reading achievement level, anthropomorphic attitudinal set, underdog attitudinal set, and culturally-alien attitudinal predisposition.

Three hundred twenty intermediate grade children were initially tested with standardized and informal instruments to obtain the information necessary to control the variables specified above. Of these, one hundred fifty were retained for the study, since the preliminary test data indicated that their intelligence levels, reading achievement levels, and attitudinal responses were distributed over the entire range of those traits.

Results indicate that the consociations between attitude and reading comprehension were not affected by race. Consequently, materials written primarily to develop or enhance reading comprehension for certain racial groups appear to offer little promise. The study also found that literal and interpretative reading comprehension are differentially affected by intelligence level, sex, and socioeconomic status. This finding raises an argument for providing different stories for various individuals representing sundry classifications on those categorical variable scales.

A STUDY OF THE REASONING ABILITY AND ACHIEVEMENT OF HEARING IMPAIRED CHILDREN WITH IMPLICATIONS FOR LINGUISTIC METHODS OF INSTRUCTION OF NORMAL CHILDREN, *Dr. James D. Beaber*, Louisiana State University; *Dr. Milton D. Jacobson*, University of Virginia and *Dr. Paul McClelland*, Virginia School for the Deaf

The purposes of this study were to evaluate the reasoning ability and verbal and non-verbal achievement of hearing impaired children to determine if significant differences in nature and degree of abilities and achievements occur between deaf and less hard of hearing children.

A sample of 24 hard of hearing and of 28 deaf students were compared. Reasoning abilities were determined by the Chicago Non-Verbal Examination and the Wechsler Adult Intelligence Scale (Performance). Achievement was measured by the verbal and non-verbal subtests of the Stanford Achievement Tests and by school grades. Other variables considered were sex, chronological age, etiology, average decibel loss, grade placement, and the hearing impairment of parents.

Computerized variance and covariance techniques were utilized to control variables found to effect relationships.

Significant differences were found. Implications for instruction of hearing impaired children and for language instruction of normal children utilizing phonics and linguistics methods were drawn.

Session 8.9 **Division C**

VERBAL LEARNING NO. 2

Robin S. Chapman, University of Wisconsin, Chairman

THE TRAINING OF MNEMONIC TECHNIQUES FOR PA LEARNING IN CHILDREN, *Mary Sue Ammon* and *William D. Rohwer, Jr.*, University of California, Berkeley

The purpose of the study was, first, to explore the possibility of training children in the use of mnemonic techniques for paired associate learning (PAL) and, second, to examine the relationship between trained performance and socioeconomic status.

One hundred twenty second-graders, half lower strata Negro and half upper strata white children, were given a pretest consisting of two separate PAL lists, one presented auditorially (auditory mode presentation) and the other both auditorially and visually (mixed mode presentation). Children were assigned to one of three treatments on the basis of their pretest scores: training, practice or control. In the training treatment condition children received instruction on the following techniques: the vocal rehearsal of both auditory and visual pairs, the visualizing of pairs presented auditorially, and the creation of imagined action sequences for both auditory pairs and pairs pictured in a stationary side-by-side arrangement. These generated action sequences were accompanied by simple sentences in which the pairs to be remembered were predicatively related (e.g., The *TEETH* bite the *APPLE*). Subjects in the practice treatment also learned to rehearse auditory items but primarily spent their time learning increasingly longer PAL lists in the traditional manner. The posttest, consisting of two PAL lists constructed in the same way as the pretest, was administered to each subject, usually during the week that followed training or practice.

Analyses, primarily analysis of variance and post hoc comparisons, revealed that the efficacy of training varied with stimulus presentation, trial, and SES variables. Development of such training techniques may have implications for improving learning efficiency in the classroom situation.

VERBAL FACILITATION OF PAIRED ASSOCIATE LEARNING AS A FUNCTION OF SYNTACTIC AND SEMANTIC RELATIONS,

Linnea C. Ehri and *William D. Rohwer, Jr.*, University of California, Berkeley

Recent investigations by Rohwer and his associates have explored various aspects of verbs which influence their ability to facilitate the paired associate learning of nouns. By examining specific functions served by verbs in their role as noun connectives, the present study aimed to show that the functions critical for facilitation involve the specification of syntactic and semantic relations.

To study these functions, the task of learning 20 pairs of nouns by a study-test method was given to 112 fourth and fifth grade children. In the experimental conditions, the study trial consisted of the presentation of each pair embedded in a sentence and connected by a verb unit. A 2x3x2 design was used to determine the effects of the following independent variables: the extent of semantic overlap of the verb with its nouns (maximal overlap with subject and minimal with object noun, called S-related verb, vs. maximal overlap with object and minimal with subject noun, called O-related verb); the character of the test-trial stimuli (subject noun vs. verb unit vs. subject noun and verb unit); and the type of verb construction (one-unit transitive verb vs. two-unit verb-preposition combination). The control group was presented with nouns linked by conjunctions.

Performance varied as a function of the interaction of Verb relation with Test stimuli. Whereas variations in test stimuli made no difference with S-related verbs, they did affect performance with O-related verbs, where no facilitation was produced by the N test stimulus but substantial amounts appeared with the other two stimuli. These results suggest that verbs, in order to facilitate learning, must be aroused during the test trial so that they can permit retrieval of the verb phrase where the object noun is stored. Contrary to a previous finding, verbs proved equal to nouns as test stimuli. Also, verb-preposition connectives were as effective as verbs, thus increasing the number of types known to facilitate learning.

PAIRED-ASSOCIATE LEARNING AS A FUNCTION OF AROUSAL AND TIME OF RECALL, *Marcia A. Lovejoy* and *Frank H. Farley*, Wisconsin Research and Development Center for Cognitive Learning

Numerous studies have dealt with the relationship between physiological activation level and the efficiency of memorizing verbal material. An experimental variable not often considered in these investigations, however, is time of recall. The present study compared learning under high and low arousal for two recall intervals. While recording galvanic skin resistance (GSR) and heart rate as measures of arousal, 32 *S*s were presented with 8 paired-associates for learning. Each *S*s eight GSR deflections to the paired-associates were then ranked. The three largest deflections were designated as high arousal learning and the three smallest deflections as low arousal learning. A similar analysis was done for heart rate, the faster heart rate indicating high arousal.

Learning under high and low arousal defined by amount of GSR deflection revealed a significant interaction ($p = .01$) between arousal level and time of recall. Learning under the low arousal condition was excellent at immediate recall but retention decreased rapidly over the next 24 hours. High arousal learning on the other hand showed low immediate recall but high 24 hour recall. There were no significant trends with high and low arousal defined by heart rate.

The GSR data would indicate that paired associates learned under low arousal exhibited high immediate recall value of rapid forgetting. High arousal paired associates exhibited a marked reminiscence effect, that is, low immediate recall and high long term recall.

AROUSAL AND RETENTION IN FREE-RECALL LEARNING, *Jacqueline E. Haveman* and *Frank H. Farley*, Wisconsin Research and Development Center for Cognitive Learning

One theory (Walker and Tarte, 1963) relating arousal to retention holds that high arousal during learning can degrade immediate retention but facilitate long-term retention relative to the effects of low arousal. The few published studies relevant to this notion have in most cases employed paired-associate (P-A) paradigms where the stimulus associated with arousal is present during both learning and recall. One method of partially avoiding this problem is the use of free-recall learning (FRL). The present research employed the FRL paradigm in two studies.

The first study employed was a 3x2 design with 48 undergraduates as *S*s, three list conditions (high arousal, low arousal, mixed) and two retention intervals (immediate and 3 days). Arousal value of the words was based on *a priori* assignments as used by Walker and Tarte. Contrary to previous P-A studies, no significant interaction between arousal and retention interval was obtained, nor was reminiscence detected. However, although arousal conditions were not differentiated on immediate retention, significantly greater *long-term retention* was associated with high- as opposed to low-arousal. The mixed-list data suggested some contextual effects.

The second study attempted to avoid the definition of arousal in terms of the material being learned, where it is difficult to control all list differences other than their "arousal category," through independent *manipulation* of arousal using white noise. A within-*S* design based on an earlier P-A study by Berlyne *et al* (1966) employed 20 undergraduates as *S*s, with one-half learning under high-arousal and the other half under low-arousal. The list consisted of 10 highly pronounceable CVC nonsense syllables with Glaze association values of 87-93 percent. Following Berlyne *et al*, performance on two training-test trials on day 1 was compared with performance on a test trial 24 hours later. The high-arousal condition was slightly *inferior* in day 1 performance but significantly *superior* ($p < .05$) in day 2 performance compared to the low-arousal condition.

These studies indicate that factors labelled "arousing" can be associated with enhanced long-term retention in FRL. Various theoretical interpretations of the results are discussed.

Session 8.10 **Division C**

CONCEPT LEARNING NO. 1

Wayne C. Frederick, University of Wisconsin, Chairman

THE EFFECT OF AVAILABLE INSTANCES UPON THE RELATIONSHIP OF MEMORY ABILITIES TO PERFORMANCE IN A CONCEPT LEARNING TASK, *Daniel D. Blaine* and *J. L. Dunham*, The University of Texas at Austin

It has been established that making past instances available in a concept learning task (CLT) facilitates performance. This effect has been interpreted as due to a reduction in the memory requirement of the CLT. If the solution of the CLT demands the recall of previously presented information, then the relationship between short-term memory abilities and performance in the CLT should be reduced by the introduction of available past instances.

Six tests representing six short-term memory abilities from Guilford's Structure of Intellect model were administered to 60 subjects prior to the administration of a CLT. The CLT was a four-category conjunctive concept problem presented in two phases. The first phase consisted of 24 concept instances, 12 of which were repeated with 1, 2, or 3 intervening instances. This provided a measure of subjects' ability to recall previously presented instances. In the second phase the availability conditions were introduced and corresponded to either zero or one previous instance available. Thirty subjects were randomly assigned to each of these conditions.

A factor analysis of the six memory tests yielded two factors—a memory for units factor and a memory for classes factor. The loadings of the repeated instances measure for the total group and trials-to-criterion for each of the availability conditions were then estimated by a matrix extension procedure. The repeated instances measure of the ability to recall previous instances loaded substantially on the memory for units factor and near zero on the memory for classes factor. The loading of trials-to-criterion on the memory for units factor was greatly reduced by the introduction of an available past instance. The loading on the memory for classes factor, however, was unaltered by the introduction of availability.

The results support the contention that a memory requirement is reduced by making past instances available. Although there are indications that additional memory abilities may be related to successful solution in concept tasks, their importance does not appear to be directly related to the recall of previously presented instances.

THE EFFECTS OF TACTUAL AND VISUAL MODES OF STIMULUS PRESENTATION OF REVERSAL AND NON-REVERSAL CONCEPT LEARNING, *Richard Bloom* and *Jane Moore*, Rutgers University

The usual paradigm for studying reversal and non-reversal learning involves the visual presentation of stimuli. The present investigation is concerned with the comparative effects on concept attainment of presenting stimuli either visually or tactually.

The experimental task consisted of a set of wooden blocks which varied along three dimensions: size, shape and texture. The blocks were constructed so that the same information is available regardless of mode of presentation. Using a counterbalancing procedure, subjects were required to correctly identify instances of one of the above dimensions as defined by a preshift criteria of ten consecutive correct classifications of positive and negative instances of the concept. Once the criterion was reached, half of the subjects were shifted to an intradimensional problem (reversal) and half to an extradimensional problem (non-reversal). For half of each shift group, the mode of presenting stimuli (either visually or tactually) remained constant for both pre and post shift problems. For the other half, the mode of presentation changed from a visual to tactual or tactual to visual presentation. Eighty children equally divided between males and females between the ages of 12-14 were randomly assigned among the eight experimental conditions. The following findings were obtained. (1) Regardless of method of presenting stimuli, the expected superiority of reversal over non-reversal learning was noted. (2) There was no significant difference in the main effect of mode of presentation for either the pre or post shift problems. (3) A significant interaction was found between type of shift and mode of presentation. In the reversal situation, a tactual presentation enhanced post shift concept attainment, whereas in the non-reversal task a tactual presentation markedly interfered with speed of attainment on the second problem. Implications of the results are related to training procedures for sightless individuals.

THE EFFECT OF VERBALIZATION OF RELEVANT AND IRRELEVANT DIMENSIONS ON CONCEPT FORMATION, *Floyd McCain, Jr.*, Indiana University; *Myron H. Dembo*, University of Southern California; *Dean H. Hustuft*, Indiana University

The purposes of this study were, (1) to elaborate on Kender and Kendler's (1961; 1962) research on the relationship between verbalization and concept formation, which has dealt with verbalization of either a relevant or irrelevant dimension, but not both simultaneously, and (2) to investigate the possible potency of different stimulus classes.

It has been hypothesized that language makes stimuli more noticeable and as a mediating response serves to orient one toward a particular stimulus dimension. The present investigators further hypothesized that if a mediating response was lengthened during initial learning to include

the second dimension of a two dimension discrimination task, this new verbalization would help the learner respond to the previously irrelevant dimension (nonreversal shift).

Eighty-four second grade Ss were randomly assigned to twelve conditions in a 2x3x2 factorial design involving two shift variables (reversal and nonreversal), three verbalization conditions (no verbalization, one dimension, and two dimension), and two discrimination variables (size and color). The Ss were tested on a discrimination apparatus with the number of trials to criterion on the final discrimination task as the dependent measure.

Significant main effects were found for reversal and nonreversal shifts ($p < .05$) and discrimination variables ($p < .05$). Although no significant interaction was found, the data indicated that during transfer learning verbalization of both dimensions facilitated the acquisition of a size discrimination to a greater extent than a brightness discrimination. (This distinction did not appear to be present in the non-verbalization condition.) The data are discussed in relationship to Kender and Kendler's mediation hypothesis and recent findings by Wolff.

THE EFFECT OF COGNITIVE STYLES IN VERBAL AND PICTORIAL CONCEPT FORMATION TASKS, *Floyd McCain, Jr.* and *Laurence Brown,* Indiana University

Cognitive style, as previously investigated by Kagan, Moss and Sigel, has been related to a number of personality and intellectual factors. Previous studies have shown that learner cognitive style interacts with concept attainment when the stimulus dimensions are pictorial. The present investigation attempted to extend this relationship to concept formation tasks involving verbal stimuli.

Forty-two college undergraduates were classified as analytic, categorical or relational according to their responses on the Sigel Cognitive Style Test and were randomly assigned to verbal or pictorial concept learning conditions. Subjects were presented a series of slides involving the paired association of nine three-letter nonsense syllables (Glaze list) with single words or pictures representing nine different concepts. Each set of slides (verbal or pictorial) included three types of concepts—analytic, categorical and relational—and three different concepts for each concept class. Four positive instances for each of the nine concepts were presented in association with the appropriate syllable under a standard paired-associate procedure. The number of errors made during the last four trials in each concept class was used as the measure of performance.

An analysis of variance failed to replicate the finding of an interaction between learner cognitive style and concept class for either the verbal or pictorial conditions. However, a significant main effect on learner style for the verbal condition revealed that analytic subjects performed better across all concept classes than those having categorical or relational styles. It thus appears that analytic college students have an advantage over those with other cognitive styles in concept learning tasks of various kinds which involve verbal stimulus dimensions.

Session 8.15 **Division D**

RESEARCH ON TEACHING

Alfred D. Garvin, University of Cincinnati, Chairman

VALIDATION RESULTS: PERFORMANCE TESTS OF TEACHING PROFICIENCY IN VOCATIONAL EDUCATION, *W. James Popham*, University of California at Los Angeles

This report describes the results of the second of two related projects designed to assess the merits of a previously untried procedure for measuring teaching proficiency. Results of the first project, conducted in the field of social studies, was reported at last year's A.E.R.A. meeting.

Two performance tests of teaching proficiency in the field of vocational education were developed during the second project, one in the field of auto mechanics (carburetion) and one in the field of electronics (power supplies). An assessment was made of each test's ability to distinguish between experienced teachers and the non-teachers with respect to these individuals' ability to achieve pre-specified instructional objectives. All subjects, teachers and non-teachers, were given sets of operationally defined objectives, and each attempted to achieve those goals during an instructional period of approximately 10 hours. Pre- and post-tests based explicitly on the objectives were given to each subject's pupils, and average class achievement was used as the index of the teacher's proficiency.

In all, 28 auto mechanics teachers and 28 non-teachers instructed over 1,200 pupils, while 16 electronics teachers and 16 non-teachers instructed over 700 pupils. Comparisons of pupil performance data revealed no systematic differences between the performance of the teacher and non-teacher group for either auto mechanics or electronics. These results were attributed to problems associated with the training of teachers as well as to the reinforcement contingencies operating when they commence teaching.

GOAL ORIENTED TEACHING EXERCISE: METHODOLOGY FOR MEASURING THE EFFECTS OF TEACHING STRATEGIES, *Russell A. Hill*, Temple University; *Donald M. Medley*, Educational Testing Service

A Goal Oriented Teaching Exercise (GOTE) is a complex methodology for relating quantifiable patterns of teaching behavior to student gains. A GOTE consists of (1) a four-day teaching unit, (2) specific educational objectives for students, (3) tests to measure student gains toward the specified objectives, (4) suggested teaching strategies in behavioral terms, (5) instrumentation and procedures for recording teacher behavior, and (6) data analysis procedures for relating student gains to patterns of student behavior.

Teachers are asked to teach a unit of work for specified educational objectives. The unit content is organized along two dimensions, kinds of information (causes, effects, and so on) and levels of abstraction (facts, methodology and generalization), to form twenty-six content cells. Educational objectives are specified for each cell. Student gains are measured during the teaching unit by equivalent forms of a test, with items related to each content cell of the unit. Teacher behavior is recorded three ways—by process, by content, and by goal. Each statement of the teacher is coded (1) using the OScAR 5V for information about the teaching process, (2) by reference to the content cells of the total unit, and (3) by educational goals.

An initial GOTE has been developed and piloted. Four teachers taught the same content unit to two classes. In one of their classes, the goal was the "recall" dimension of Bloom's Cognitive Taxonomy; in the other class, teaching was focused on "application" of knowledge. Analysis of the data indicated that student gains differ both on the goal dimension and by different teachers. Teacher behavior data also distinguished between teachers, with twenty-nine of the sixty-four OScAR keys showing significant differences among teachers. There appears to be an interaction between measures of student gains, patterns of teacher behavior, and goal dimensions. Problems encountered and plans for future development of the methodology are outlined.

TEACHER ASSESSMENT AS A PROBLEM OF SAMPLING WITHIN A COMPLEX COMPOSITE CRITERION, *K. B. Start*, University of Saskatchewan, Regina Campus

Some of the work on teacher effectiveness has produced positive results, some negative, but most have been equivocal. Many criteria of teacher effectiveness have been used and appear to be little related to each other. Most studies in this area have, to a greater or lesser extent, foundered on the criterion issue.

A series of studies in England indicated that judgments of a teacher's effectiveness were based on a multiplicity of teaching roles. Different types of schools based their estimates on different samples of these roles. This sampling reflected the philosophy and values from the school type. By taking a group of mythical teachers, assigning a series of random scores for these roles for each of the teachers, and then using various sampling strategies, the lack of agreement between the overt criteria can be vividly demonstrated. The correlations found in practice in the U. K. studies can be, and were, replicated by means of this sampling approach.

The thesis then is that the lack of agreement between single estimates of teacher efficiency is due to the latter being a multiple measure, and the overall single assessment is actually a composite derived from the scores made on the basis of sampling from within these many subcriteria. The sample taken depends on the values of the assessor, and hence different judges having different values will take different samples and would have composite scores which differ as a function of their value dissonance.

THE RELATIONSHIP OF STUDENT TEACHERS' OPINIONS OF INTELLECTUAL EMPHASIS AND JUDGED COGNITIVE LEVELS OF QUESTIONS FRAMED FOR DISCUSSION AND TESTS, *Drew C. Tinsley*, University of Missouri at St. Louis; *O. L. Davis, Jr.*, The University of Texas at Austin

This study, part of a larger investigation of the cognitive levels of questions secondary social studies student teachers planned to guide discussions and to use as tests, examined the relationship between the judged cognitive levels and the student teachers' opinions of the intellectual emphasis of the questions they framed.

Ss were 67 social studies student teachers at a southwestern university with at least six weeks student teaching experience. Assigned randomly to one of four experimental groups according to grade level (8th or 11th) and instructional task (discussion or test), Ss planned questions based on the same stimulus material, a digest of congressional testimony. Immediately following their completion of the experimental task, Ss indicated the intellectual emphasis intended for the set of questions for each of six descriptive categories (memory, convergent thinking, logical thinking, reflective thinking, divergent thinking, and evaluation). Ss marked the intended emphasis on a graphic rating scale, and separate scores for each scale were derived. Ss' questions were rated according to their cognitive level by three trained judges. Inter-rater agreement exceeded 90 per cent. Obtained data were treated by correlational analyses.

Results revealed that student teachers' opinions correlated nonsignificantly, for the most part, with their categorized questions. Significant negative correlations were obtained between student teachers' opinions about their emphasis on reflective thought and with their questions categorized as convergent and evaluative. Their opinion of emphasis on evaluation was significantly associated with the number of their questions rated as evaluation. The data indicate rather clearly that these student teachers were unable to predict with accuracy the intellectual emphases of their questioning. These findings are discussed in terms of specific suggestions for teacher education components and evaluation procedures in teacher education.

Session 9.1 **Division C**

PRESCHOOL TRAINING PROGRAMS NO. 2

Miles Friedman, University of South Carolina, Chairman

THE PRE-SCHOOLER'S RELATIONAL CONCEPTS AS TESTED IN OBJECT AND REPRESENTATIONAL FORM, *Selma Greenberg* and *Ruth Formanek*, Hofstra University

The purpose of this study was to find a normative base for the construction of new pre-school curricula of particular relevance for those labeled "disadvantaged." Our testing was conducted in both middle-class nursery schools and Economic Opportunity Day Care Centers where lower-class children are served. Two hundred and one children composed our sample and the following tests were administered: a relational concept test in object form (based on Boehme's test), a relational concept test in representational form, the Caldwell Pre-school Inventory, the Harris Test for Lateral Dominance. Our set of hypotheses included: (a) differences in performance according to social class would be found, (b) a hierarchy of difficulty in concepts exists, (c) scores on representational concepts and preschool inventory would be positively correlated, (d) relational tasks in object form would be simpler to solve, especially for lower-class youngsters, than similar concepts in representational form, (e) children with established lateral dominance would be more successful than those with mixed dominance. The results confirmed hypotheses a, b, and c, disproved hypothesis e, and were insufficient to either completely confirm or disprove hypothesis d.

A surprising finding was the results of one O.E.O. center whose students achieved close to the norms achieved in one of the middle-class nursery schools. Our subjective evaluation of the center's program led us to suspect that, if our findings were not spurious, they might be attributed to the more structured and concept-oriented curriculum of that particular center. In deciding to investigate this issue more closely, a curriculum in concept achievement in six areas, including that of relational concepts, was constructed. This curriculum will be tried in a controlled experiment in the O.E.O. center whose children has the lowest norms in our original study.

COMPARATIVE EFFECTIVENESS OF SPEAKING VERSUS LISTENING IN IMPROVING THE SPOKEN LANGUAGE OF DISADVANTAGED YOUNG CHILDREN, *Willa Gupta* and *Carolyn Stern*, University of California, Los Angeles

The language of the disadvantaged child is deficient in those syntactical features which underlie logical operations and conceptual thought. Remediation of this critical deficit requires a more intensive and structured program than the experience-approach which characterizes the middle-class nursery. However, there has been little controlled research to examine specific aspects of structured programs to determine whether overt verbalization plays an important role in this learning.

The major hypothesis, that children who repeat sentences aloud will acquire greater facility in forming similar sentences on their own than those who only listen to the spoken sentences, was tested with 40 Negro children, 43 to 55 months old. The identical sequences of five 15-minute daily lessons plus 2 days of testing were presented under 2 treatment conditions, speaking and not-speaking. Each lesson contained approximately 15 problems. A problem consisted of a sequence of 5 pictures and taped commentary such that the first 3 frames presented the identical sentence structure, varying only the action verb. The fourth frame was a 3-choice selection in response to 1 of the 3 sentences. On the fifth frame the child either listened to or produced the sentence to 1 of the first 3 frames. The treatment difference consisted solely in whether the child listened to the sentence repeated twice or listened and repeated the sentence. Of the total posttest, which tested both selection and verbalization, analysis of covariance (with pretest as covariate) showed a significant difference (.01 level) favoring the speaking group. When the scores were analyzed separately, no significant treatment effects for the selection task were found. The significant difference was attributable to the scores on verbalization. Similarly, a transfer test with verbalization to new pictures produced significant treatment effects ($p < .01$).

POSITIVE EFFECTS OF A BICULTURAL PRESCHOOL PROGRAM ON THE INTELLECTUAL PERFORMANCE OF MEXICAN-AMERICAN CHILDREN, *Ronald Henderson*, University of Arizona; *Richard Rankin*, University of Oregon; *Mary Frobisher*, Director, Bicultural Nursery School, Tucson, Arizona

The purpose of this investigation was to test the assumption that the intellectual performance of disadvantaged Mexican-American children may be favorably affected by integrating these children into classes with Anglo peers from more advantaged backgrounds in a Bicultural Nursery Program (BNP).

An experimental group of 18 Mexican-American five year old *S*s were matched by initial WPPSI score with 18 Mexican-American *S*s in a conventional Head Start program (HSP) and with 18 non-program Mexican-American *S*s (NP) living in poverty areas of Tucson, Arizona.

The interval of the study was 9 months and the pre and post test measures consisted of the Wechsler Preschool and Primary Scale of Intelligence.

Conventional matched group t tests indicate that all groups make some gain over time. The gain of the BNP group exceeds the NP control group by 7.11 total I.Q. points ($p < .02$) and exceeds the gain of the HSP group by 5.55 points ($p < .05$). The gain of the HSP group does not exceed the gain of the NP control. The bulk of the BNP gain comes from an 11.22 Verbal I.Q. gain over the NP control. There is no net gain for BPN or HSP over NP in Performance I.Q.

Since the net verbal I.Q. gain with the BNP group equals their standard deviation in initial verbal score and the within group variance decreases between tests, there is a strong suggestion that a bicultural program in which a relatively small proportion of disadvantaged are mixed with a larger proportion of advantaged children may induce permanent intellectual change.

PREKINDERGARTEN PROGRAMS FOR EDUCATIONALLY DISADVANTAGED CHILDREN: THIRD-YEAR EVALUATION AND FOLLOW-UP, *Ruth Salter* and *Louis T. Di Lorenzo*, New York State Education Department

The study involves three successive waves of approximately 700 four-year olds who have been pretested on the Stanford-Binet and the PPVT and randomly assigned to experimental and control groups in eight school districts. Both those children attending prekindergarten and those remaining at home are being followed into second grade.

The report will include (1) an analysis of pre-post I.Q. and language scores for the third and final wave of project children and a comparison of their results with those of the first two waves, (2) an analysis and comparison of readiness test scores at the end of kindergarten for the first two waves of children, and (3) an analysis of achievement test results at the end of first grade for that group of Wave I children in which there were significant differences between experimentals and controls at the end of prekindergarten. Sex, race, socioeconomic status, and type of program will be considered in the various analyses.

The prekindergarten programs of the eight school districts represent divergent points of view in preschool education. Several are traditional nursery school programs; one has employed a teaching machine in an academically oriented program; and another is using an adaptation of the language pattern drills developed by Bereiter and Engelmann. Two of the programs have heterogeneous classes of disadvantaged and non-disadvantaged children; the others are homogeneous. All have as their goal the improvement of capacity to learn, language development, self-concept, physical development, and attitude toward school.

TEACHING CLASSIFICATION CONCEPTS TO DISADVANTAGED PRESCHOOL CHILDREN, *Anne M. Bussis, Jean Orost,* Educational Testing Service; *Miriam Papanek,* Sarah Lawrence College; *Masako Tanaka,* Educational Testing Service

This pilot study was designed to investigate whether the classification competence of disadvantaged kindergarten children could be improved by instructional activities taught by a regular public school teacher and designed to fit into the normal classroom routine. In addition to these practical issues, the study was also designed to investigate the relationship of impulsivity to categorizing behavior.

Children in two morning kindergarten classes in a Brooklyn school received instruction, while the afternoon classes of the same two teachers served as a comparison group. Each child was given an individually administered pretest and posttest consisting of a specially developed classification measure and the Draw-A-Line-Slowly (DALS) test which measures impulsivity. Following pretesting, the teachers presented instructional activities in classification three times a week over a four week period. Major results were as follows:

(1) Children in the two experimental classes made significant gains ($p = .05$) on the classification measure, while children in the comparison group did not.

(2) There was a significant correlation in the experimental group between *initial time scores* on the DALS and *gain scores* on the classification measure ($r_s = .41$). This same correlation for the comparison group was *—.06*. Initial time scores and initial classification scores were essentially unrelated for both experimental and comparison groups, however ($r_s = .03$ and *—.02* respectively). These data suggest that impulse control and classification skill may not be related in any "natural" developmental way among disadvantaged children, but that impulse control becomes a crucial factor in the child's ability to benefit from a typical kind of classroom instruction.

(3) Children high on impulse control improved significantly more than their peers on problems requiring the separation of form from meaning and on problems in which manipulation of objects was minimized. There was no significant difference between impulsive and nonimpulsive children's performance on other types of problems.

Session 9.2 **Division C**

MICRO-TEACHING

Michael Orme, Indiana University

EFFECTS ON THE VERBAL TEACHING BEHAVIORS OF BEGINNING SECONDARY TEACHER CANDIDATES' PARTICIPATION IN A PROGRAM OF LABORATORY TEACHING, *O. L. Davis, Jr.,* The University of Texas at Austin; *B. R. Smoot,* University of Arkansas

This study was designed to assess the effects of a Teaching Laboratory (TL) component on the verbal teaching behaviors of beginning secondary teacher candidates. The TL, based on micro-teaching procedures, is integral to the regular introductory course in secondary school teaching in a teacher education program. In the TL, candidates teach scaled-down lessons focusing on specific pedagogical tasks in a teach- re-teach cycle. Systematic feedback from pupils, instructor, and audio-recordings is available to the candidate. Pupils in the TL are peers.

Ss were 140 beginning secondary candidates enrolled in six sections of the first course in teaching in the professional sequence at The University of Texas at Austin. Ss in Group A (N = 85) were enrolled in sections incorporating the TL component. Ss in Group B (N = 55) were enrolled in sections which did not incorporate the TL. During the experimental period, Ss in Group A taught four TL lessons whereas Ss in Group B taught none. All Ss taught a ten-minute "pre-test." Following the seven week experimental period, all Ss taught a ten-minute "post-test." The criterion measure was the *Laboratory Observation Schedule and Record (LOScAR)*, a modification of *OScAR 5V* developed by Medley and others (1968).

Twenty-two variables, consisting of the 13 *LOScAR* category scores and nine ratio scores, were analyzed by ANCOVA procedures. Results revealed significant differences ($p < .01$) between TL and non-TL groups on 18 of the variables. All differences were interpreted as favoring Ss with TL experience.

Findings are discussed with respect to the effectiveness and utility of the TL in regular teacher education programs and to the value of specific laboratory experiences prior to student teaching as means of enabling teacher candidates to become competent in interactive behaviors.

STUDENT ATTITUDES AND ACHIEVEMENT IN AN EDUCATIONAL PSYCHOLOGY COURSE AFTER MICRO-TEACHING, *Adrian P. Van Mondfrans, Thomas M. Smith,* and *John F. Feldhusen,* Purdue University

The purpose of this study was to measure the effects of micro-teaching on student attitudes and achievement in an undergraduate educational psychology (EP) course.

Students in an EP course met two days a week for lecture, discussion, and TV presentations and then were divided randomly into five groups for separate meetings on Fridays. One group received the micro-teaching treatment (MT). The other four groups discussed the course reading material, unit tests, and assignments.

Subjects in the MT group prepared an 8-minute lesson and presented it to their peers. The presentation was video-taped. *Ss* viewed the taped performance, discussed it with *E*, and retaught the same lesson. When not involved in MT teaching, the *Ss* either acted as the audience or heard a short lecture on classroom management or presentation skills by *E*.

Criterion measures were scores on unit tests and papers, the Purdue Rating Scale for Instruction (PRSI), an attitude measure, and ratings of the micro-teaching experience.

Scores on unit tests and papers showed no significant difference between the MT and other groups, although *Ss* in MT did not receive help for tests or papers similar to that in other discussion groups. Three PRSI items which mentioned lab facilities or asked how the course met *S* s needs showed significantly higher ratings by the MT group. On PRSI items pertaining to other aspects of the EP course the ratings of the MT group were lower than those of the other discussion groups (of 23 *F*-ratios 16 were significant at $p < .05$). When all 5 groups were asked how much they had learned in their discussion groups, the MT group responded with significantly higher estimates. When *S*s estimated the amount of teaching-relevant material they had learned in the discussion sections, the difference was even greater.

THE EFFECTIVENESS OF MICRO-TEACHING AND SELF-INSTRUCTIONAL MODELS IN THE ACQUISITION OF TWO TEACHING BEHAVIORS SIMULTANEOUSLY: MAINTAINING PUPIL-TASK ORIENTATION AND LECTURING, *David B. Young*, University of Maryland

The purpose of this study was to appraise the relative effectiveness of various modes of modelling for training teachers in two teaching behaviors simultaneously without direct supervisor-teacher conferences (self-instructional).

Interns taught three five-minute lessons to five students who role-played six different non-task oriented behaviors. The various modelling protocols were presented between teaching sessions.

The lecturing skill, repetition, was modeled on video-tape with two variations: (1) subjects were focused contingently (comments and visual prompts recorded on the second sound track parallel to the behavior to be learned) and (2) non-contingently by a written guide.

Two modelling protocols for maintaining pupil-task orientation were used. The first was a sixteen millimeter film presenting eleven "disciplinary" incidents with three alternative teaching responses to each. This film emphasized *desistance* techniques. The second was a video-tape model (with contingent focus) emphasizing control techniques in *prompted reward* and *withheld sanctions* categories (reinforcing instructional responses and ignoring deviant behavior) and a symbolic model emphasizing all three categories.

The following results were obtained:

1. The second occurrence of a deviant behavior within a five-minute class period stimulated a greater number of different responses than the first occurrence.
2. Subjects viewing a combination of models acquired a larger repertory of alternative control techniques on certain variables.
3. Subjects viewing the video-tape model and studying the symbolic model used a greater number of *different* techniques on certain variables and tended to use more alternative responses on each successive teaching episode.
4. Subjects exhibited limited improvement on both lecturing and discipline variables but failed to reach significant levels.
5. Subjects on all groups used a greater number of *withheld sanctions* and fewer *desistance* techniques.

THE EFFECTIVENESS OF INDIVIDUALLY PRESCRIBED MICRO-TEACHING TRAINING MODULES ON AN INTERN'S SUBSEQUENT CLASSROOM PERFORMANCE, *Dorothy A. Young*, Johns Hopkins University; *David B. Young*, University of Maryland

The purpose of this study is to assess the effectiveness of individually prescribed micro-teaching training modules in the acquisition of selected teaching behaviors (pre-internship) and subsequent implementation of these behaviors in classroom teaching (internship).

Two groups of twenty interns comprise the experimental groups. The control group begins their intern teaching assignment without the pre-internship micro-teaching program. The experimental group taught a diagnostic lesson at the beginning of the semester preceding their internship. This performance was analyzed using the Medley and Mitzel Observation Scale and Record and Hough's Observation System for the Analysis of Classroom Instruction. Based upon this assessment, a sequence of approximately four micro-teaching training modules was individually prescribed for acquiring specified teaching behaviors. Midway in the same semester, the Oscar and Hough instruments were administered again and another sequence of micro-teaching modules prescribed.

The data are collected by administering the Oscar and Hough Observation System for Instructional Analysis on each intern's classroom teaching performance during the fourth or fifth week and again during the ninth or tenth week of the internship.

The data are analyzed by comparing the mean frequency and number of different specific teaching behaviors and patterns in a twenty-minute sample. Preliminary evaluation of the data indicates that the experimental group has acquired a greater number of specific teacher behaviors and teaching patterns.

Session 9.3 **Division C**

CREATIVITY NO. 1

Hilliard Jason, Michigan State University, Chairman

A TEST OF THE McNEMAR HYPOTHESIS ON THE RELATIONSHIP BETWEEN CREATIVITY AND INTELLIGENCE IN TWO SEPARATE CULTURES, *Anthony J. Allen, John S. Dacey,* and *George F. Madaus,* Boston College

The purpose of this study was to test McNemar's hypothesis that ". . . at the high IQ levels there will be a very wide range of creativity, whereas as we go down to the average IQ and down to the lower levels the scatter for creativity will be less and less."

The sample consisted of 867 American eighth grade students from twenty-three junior high schools in three states and 182 Irish secondary school students from suburban Dublin. In the American sample divergent thinking was measured by four tests adopted from those of Torrance. The *Lorge-Thorndike* was used as a measure of intelligence. Selected tests from the *Minnesota Tests of Creative Thinking* and the *AH4 General Test of Intelligence* were administered to the Irish sample. In both samples the IQ distributions were divided into three equal ranges. A one-way analysis of variance and a correlational analysis were performed.

Results offered only slight support for the McNemar hypothesis. The authors suggest three possible explanations of these findings. Their major hypothesis has to do with a number of validity problems associated with creativity tests.

THE RELATIONSHIP OF CREATIVITY, MOTIVATION AND SOCIAL CLASS: A MULTIVARIATE STUDY, *William E. Connett* and *Judy Lyon*, Educational Research Bureau, Southern Illinois University

The purpose of this study was to explore the relationship of three contextual conditions, test anxiety, and social class to creative performance.

Hypotheses tested were: (1) social class is negatively related to creative performance in a monetary incentive (MI) condition but is positively related to creative performance in a social incentive (SI) condition, and (2) test anxiety is negatively related to creativity in an SI situation and in an MI situation but has a positive relationship in a relatively neutral incentive (NI) situation. Similar results were hypothesized for both verbal and nonverbal creative performances.

After acquiring measures of subjects' test anxiety and SES, Guilford's Product Improvement Test and Torrance's Picture Completion Test were administered to three groups of 30 seventh grade subjects. All groups were instructed to give clever and unusual ideas. In addition, Group I (SI) was given the tests by a distinguished-looking man who said that it was very important that they do well on the tests because the scores would be available to their teachers and their parents and would be placed in their permanent file, Group II (NI) was given the tests by a graduate student who said that he did not particularly care how well the subjects did on the test, Subjects in Group III (MI) were given the tests by a graduate student who said he did not particularly care how well the subjects did on the tests but that he would pay them money for clever and unusual ideas, pointing out that it was possible for them to make up to five dollars.

Multiple regression analysis was used to test the stated hypotheses.

UTILIZATION OF DIVERGENT AND CONVERGENT THINKING AND PERSONALITY VARIABLES FOR THE PREDICTION OF ACADEMIC ACHIEVEMENT OVER A FOUR-YEAR PERIOD, *Robert Elias, John Feldhusen,* and *Donald Treffinger*, Purdue University

The purpose of this research was to determine (1) the level of multiple correlation between an optimum set of divergent and convergent thinking and personality variables assessed in the eighth grade and academic achievement assessed four years later, (2) the significant predictor variables, and (3) differences between males and females.

Divergent thinking was assessed with Guilford's Alternate Uses and Consequences Tests, convergent thinking with SCAT and STEP, and personality with Sarason's general anxiety scale and a self-rating scale of creativity. The criteria were STEP scores assessed four years later.

The sample, 113 males and 92 females, was gathered in 1962 and 1963 from a small city school system. The 1962 *S*s were used in validation analyses using a weighted regression analysis with a step-wise tear-down procedure. Coefficients for the variables in reduced sets were used to predict achievement for 1963 *S*s.

The multiple Rs for reduced sets in the validation analyses were all significant ($P < .01$), and for STEP mathematics, science, social studies, and reading, respectively, were as follows for boys: .66, .62, .65, and 61. For girls they were .67, .65, .84, and .71. Several cross validation Rs were not significant.

STEP reading, SCAT quantitative, originality, and creativity self-rating appeared in reduced sets for all four achievement criteria for males. The STEP, SCAT, and flexibility scores were significant predictors of all criteria for girls. Substantial sex differences were also found.

The implications of these findings seem to be that (1) academic achievement can be predicted quite accurately over a four-year period with divergent and convergent thinking and personality variables, (2) sex differences must be taken into account, and (3) self-ratings of creativity add significantly to the predictions.

THE RELATIONSHIP BETWEEN ACADEMIC GRADES AND CREATIVITY TEST SCORES DERIVED FROM FOUR DIFFERENT METHODS OF TESTING, *Donald R. Ferris, John F. Feldhusen,* and *Adrian P. Van Mondfrans,* Purdue University

This research investigated the relationship between academic grades and scores obtained by four methods of administration of the Torrance Tests of Creative Thinking (TTCT).

The TTCT were administered to four classes of fifth grade students (N = 105) and to four classes of eighth grade students (N = 112) using four different methods: (1) The Wallach and Kogan method (WK) induced playfulness and relaxation in a 20-minute session prior to testing; (2) Incubation (INC) was induced in a meeting four days before testing in which creative thinking was discussed, and students were given notebooks in which to write clever or unusual ideas; (3) Ss were given TTCT booklets to take home (TH) for four days and told to work on the exercises when they wished; (4) TTCT were given with standard test instructions (ST).

Simple correlations were calculated within each testing method among IQ, three verbal and four figural TTCT scores, and academic achievement (AA) in social studies, mathematics, English, and science. Correlations between IQ and AA were all significant (r's ranged from .39 to .75, $p < .01$) and two of 28 r's between TTCT scores and IQ were significant at $p < .05$. Methods of TTCT testing resulted in different patterns of correlation between TTCT and AA. Significant correlations ($p < .01$) between TTCT verbal scores and AA occurred as follows: (WK) 2 of 12 r's; (ST) 2 of 12 r's; (TH) 6 of 12 r's; (INC) 11 of 12 r's. Correlations between TTCT figural scores and AA yielded r's ($p < .01$) as follows: (WK) no significant r's; (ST) 8 of 16 r's significant in a negative direction; (TH) 4 of 16 r's significant in a positive direction; and (INC) no significant r's. Several of these correlations changed when IQ was partialed out.

Session 9.4 Division D

APPROACHES TO RESEARCH DESIGN

Desmond Cook, The Ohio State University, Chairman

AN EXAMINATION OF EXPERIMENTER BIAS IN BEHAVIORAL RESEARCH, *John J. Kennedy,* University of Tennessee

Two separate studies were conducted to assess the effects of experimenter (*E*) expectations of experimental outcome. Different methods of inducing *E* outcome expectancy were used in each study. Both studies, however, employed a standard "Taffel-type" verbal conditioning task, a prototype of typical interpersonal research situations, as the vehicle with which to examine *E* expectancy effects.

Experiment 1. Six *E*s were used, two who were led to expect that subjects (*S*s) would condition, two who were led to believe that conditioning would not result in the experiment, and two *E*s who were not subjected to an outcome expectation indoctrination. Two modes of presenting stimulus materials were employed—in a face-to-face situation and through the use of an overhead projector—to help assess the nature of *E* cues (auditory, facial, gestural, etc.) used to communicate outcome expectation. Ss (N = 60) were administered the conditioning task, and resultant data were analyzed in a 3 x 2 x 2 mixed design. No significant differences either among *E* expectancy groups or between stimulus presentation modes were found.

Experiment 2. Eight *E*s, two "humanistically oriented" females, two "humanistic" males, two "deterministic" females, and two "deterministic" males, were erroneously led to expect that only *half* of their *S*s *would* condition. Each *E* ran 10 *S*s (N = 80). The analysis of data in a 2 x 2 x 2 x 2 mixed design indicated that *S*s' conditioning performance conformed to the *a priori* expectations of *E*s ($p < .05$). The interaction of *E*s' philosophical orientation and sex was also found ($p < .05$) to have affected *S*s' conditionability.

The combined findings suggest that *E* bias is, in itself, a complex variable. Pre-experimental *E* expectancy was shown to have influenced interpersonal research findings; however, its effect appears to be related to the manner in which the expectation was acquired. Methodological implications of *E* bias and *E* characteristics are discussed in relation to the internal and external validity of behavioral experimentation.

RESEARCH DESIGNS AND EVALUATION RESEARCH, *Gordon A. Welty*, Pittsburgh Board of Public Education

Rigorous research designs are necessary for unequivocal research findings. Feedback of evaluative reports during the life of a program are necessary for program stability and improvement.

It has been argued by Daniel Stufflebeam, *et al*, that the two are methodologically incompatible. If this is true then current attempts to develop taxonomies of evaluation to replace experimental designs are necessary. If classical designs and evaluative feedback are incompatible, one or the other must be abandoned.

A careful examination of the situation indicates, however, that Stufflebeam, *et al*, are incorrect. We prove that, at *least* for a limiting case argument, the two are compatible. The rational manager can respond to evaluative feedback by modifying the treatment so as to guarantee program stability. The sufficient condition for compatibility of design and feedback here is that the deviance of the process p from its stable state does not exceed the experimental error of the classical design. The deviance is minimized, within the constraints of the experimental design, by negative (entropic) feedback.

More important, the manager can respond to evaluative feedback by improving the program. By the reaction R, describing the rational manager's behavior, it is shown that a change in process from p to p* during the evaluation cycle, by the equation

$$p^* = R(p)$$

is sufficient for compatibility where (p, d) is a metric space, and the contractive mapping theorem implies a unique value of p, call it π, viz.,

$$\pi^* = \pi.$$

This fixed point is the stable, improved value of p, where π can be significantly different from the initial process state and guarantees compatibility. This is a case of deviance enhancing (antientropic) feedback control, again within the constraints of the experimental designs. An examination of Fisher's *Design of Experiments* provides evidence that this proof is in accord with the classical conception of research designs. Thus, the arguments of Stufflebeam, *et al,* are misplaced, as is the need for new taxonomies of evaluation.

THE PROBLEM OF MEASUREMENT ACCURACY IN EDUCATIONAL RESEARCH, *John Klingensmith* and *John Menne,* Iowa State University

This study was undertaken to find evidence for or against the authors' contention that educational research results are often rendered ambiguous by the failure of researchers to consider the accuracy of their measuring instruments. Eleven journals reporting educational research were examined for the past five years. Data were collected for the number of research vs. non-research articles. For the research articles, data were collected for articles in which measurement accuracy was not pertinent, for articles in which measurement accuracy was considered (and whether the method of estimation used was current or obsolete), and for articles in which measurement accuracy should have been but was not considered. These data are presented in tabular form.

A typical article from each of the last two categories is analyzed in detail to point up the need for measurement accuracy and to show how its absence can render published results ambiguous if not meaningless. An article is also cited in which the requisite attention to measurement accuracy produces clear and provocative results. Finally, the implications of these somewhat dismal data are discussed, and suggestions are made of the need to bring a greater awareness of the problem to researchers now in the field and especially to those concerned with the training of future educational researchers.

GROUPS AND EXPERIMENTAL UNITS IN EDUCATIONAL RESEARCH, *David J. Wright,* Michigan State University

The choice of an appropriate experimental unit is a continuing and often misunderstood problem in areas such as classroom and counseling research, where treatments are administered in a group setting. There has been considerable discussion about losing valuable degrees of freedom and having to analyze group means when the individual is really the object of interest. Such discussions tend to overlook an important consideration, for an appropriate analysis often makes the question irrelevant.

In most designs used for experiments with groups, it is necessary to nest groups within treatments to avoid confounding treatment effects with those of the teacher, counselor, and others. Unless this replication factor is considered fixed, the subjects within groups effect is never appropriate for testing the treatment main effect and is seldom so for other effects of interest. Thus, the question of whether or not the subject can be used as the unit of analysis is generally not relevant for designs of this type.

Several common designs for use with groups in classroom and counseling experiments are outlined, and the conditions under which individual scores are not relevant to the analysis are discussed.

Session 10.2 **Division A**

QUANTITATIVE TECHNIQUES FOR EDUCATIONAL PLANNING AND EVALUATION

John Sokol, U.S. Office of Education, Chairman

AN OPERATIONAL APPLICATION OF THE STUFFLEBEAM-GUBA CIPP MODEL FOR EVALUATION, *Robert S. Randall*, Southwest Educational Development Laboratory, Austin, Texas

A program planning and evaluation system, based on the CIPP Model, is described. It involves collecting, analyzing, interpreting, and reporting *relevant information* to appropriate levels of *decision-making* at *critical times* in order *to effect* better *planning* and *operation.*

Relevant information is information that promotes and services important decisions. At least four classes of decisions are made in an organization. These are long-range planning, structuring, modifying, and recycling decisions. The system must alert decision makers to impending decisions and furnish information needed to serve each class of decisions.

Decision-making occurs not only in different classes but also at different levels of complexity and authority. In order for relevant information to promote and serve decisions at appropriate levels of complexity, the information must be reported in a manner that is consistent with the complexity of the decision level so that it can be utilized efficiently and effectively. Therefore, the nature and detail of the reported information must vary as the level of complexity of the decision varies.

The *critical time* when decisions of the several classes and levels must be made demands that reporting of information be carefully scheduled in order to service the decisions properly.

To the extent that the system alerts decision makers in the organization to the nature of some of the decisions they face and services these decisions with the information they need when they need it, then the system does function operationally to improve planning and operation.

Flow charts, graphs, and illustrations are used to clarify the operation of the model, and results (positive and negative) are discussed.

AN OVERVIEW OF OPTIMAL CONTROL THEORY APPLIED TO EDUCATIONAL PLANNING, *George S. Tracz*, The Ontario Institute for Studies in Education

Much research has been directed lately to the application of quantitative techniques to develop a number of planning models in education. A brief outline of these techniques is given and examples of these models in different areas in the field are listed. The type of educational problems that these techniques can solve is discussed. Then it is shown that a class of problems, which embodies the decision-making process followed by administrators, can be studied better by incorporating a mathematical technique known as Optimal Control Theory.

Optimal Control Theory has received much of its impetus through research carried out on Space System Analysis and Military Applications problems. However, Optimal Control Theory goes much beyond the so-called Systems Approach in which component techniques (well-known contributors in parentheses) such as Operations Research (Sisson), Systems Analysis (Stroller), and PPBS (Hartley) have made such an impact on the education scene. In fact, these techniques have created a dialogue between systems analysts and school administrators which culminated in the USOE Symposium on System Analysis and Education, November 1967.

This paper will outline the structure of the Optimal Control Theory problem formulation, will focus on the features which (theoretically) make it a most powerful technique, and will list the pioneering contributions of this approach to educational planning problems. Then, a series of problems in educational planning will be formulated in this new format. It will be shown that there exists a natural, systematic incorporation into the decision-making process in educational planning of CONTROL, identified, for example, as resources, facilities, and political constraints, and OPTIMIZATION, distinguished here as a "best" selection of *dynamically* evolving objectives and figures-of-merit. Data requirements and numerical implementation will be discussed.

THE USE OF MATHEMATICAL PROGRAMMING MODELS TO OPTIMIZE VARIOUS OBJECTIVE FUNCTIONS OF FOUNDATION TYPE STATE SUPPORT PROGRAMS, *James E. Bruno*, U.C.L.A.

For many years, the tools and techniques of operations research have been successfully applied to various resource allocation problems in business and industry. The purpose of this study was to explore the possibility of applying one such tool of operations research, namely mathematical programming, to the problem of the allocation of state resources to local school districts in foundation type state support programs.

A generalized mathematical programming model was developed to represent or simulate the fiscal aspects of foundation type state support programs. The generalized model was then evaluated for reliability and validity by applying it to the California junior college state support program. In order to insure consistency in the model, all coefficients and parameters were calculated directly from basic input data (64-65) using the current "formula" for state support to these junior college districts. This specific model was then solved under various objectives functions. These included tax rate minimization, foundation level maximization, and state costs minimization. Two major overall policy objectives were also examined in the study. The first policy objective (referred to as simple) was a state support program which distributed state funds in a manner which equalized the final total district expenditure per ADA among the junior college districts in the system. The second policy objective (referred to as complex) distributed state funds in a manner which permitted a specified percentage spread to be a variable in the model and assume whatever value was necessary in order to maximize or minimize the given objective function.

The usefulness of mathematical programming models as a method for distributing state funds to local school districts was demonstrated by the resulting non-trivial optimal solutions and the general increase in state aid to the lower local ability districts in the state support program. In one particular solution, where equalization of final total district expenditure per ADA was the major overall policy objective, 88% of the districts received more state aid with no increase in the total amount of state funds available or the tax rate.

Solutions obtained by the model, in conjunction with the ability of the model to incorporate future refinements in school finance, strongly suggest the possibility of using the mathematical programming model developed in this study, or similar models, as a method for distributing state funds to local school districts.

Session 10.3

Division B

STUDIES OF PRESCHOOL PROGRAMS

Chairman To Be Announced

A TWO-YEAR LANGUAGE ARTS PROGRAM FOR PRE-FIRST GRADE CHILDREN: FIRST YEAR REPORT, *Dolores Durkin*, University of Illinois

The purpose of this research, begun in September, 1967, is to develop and evaluate a pre-first grade curriculum based on earlier findings about preschoolers who learned to read at home. Because the common age for them to begin reading was four years, this study started with a group of 37 four-year-olds and will continue with them through the kindergarten year. The children are in two classrooms, each with a teacher and teaching assistant. Essentially, the program is language arts oriented, with a stress on a writing and spelling approach to reading which is personal and also devoid of observable pressure on the children. Connected with it is a parent education program.

Pre-experimental data include results from (a) the Stanford-Binet Intelligence Scale, (b) the Illinois Test of Psycholinguistic Abilities, and (c) the achievement tests for reading and numeral and letter identification. End-of-year testing dealt with the three types of achievement. Other sources of data include two interviews with each parent, teacher and parent questionnaires about the subjects, and classroom observations.

Socioeconomic data collected in home interviews showed that 22 of the 37 subjects were from upper-lower class families. Twelve were classified as lower-middle and 3 as upper-middle. Other data indicated that IQ's ranged from 92 to 146, with a median of 109. At the time the program began, subjects' CA's ranged from 3 years 9 months to 4 years 10 months, and their MA's from 4 years to 5 years 11 months. The correlation coefficient for CA and MA was +.27.

Results from the three achievement tests administered at the end of the year showed the three scores to be highly correlated (in the +.80's). Correlation coefficients for MA and the three achievement scores were the next highest (in the +.40's). Lowest correlations (all less than

+.10) were for CA and achievement. When achievement scores were examined in relation to the two classrooms, they were found to be significantly different. The higher scores were in the classroom in which IQ's were lower, but the teaching was obviously better.

A LONGITUDINAL STUDY OF DISADVANTAGED CHILDREN WHO PARTICIPATED IN THREE DIFFERENT PRESCHOOL PROGRAMS: TRADITIONAL, DIRECT VERBAL, AND AMELIORATION OF LEARNING DEFICITS, *Merle B. Karnes, James A. Teska* and *Audrey S. Hodgins*, Institute for Research on Exceptional Children, University of Illinois

In 1965, with supporting funds from the U.S. Office of Education, comparable groups of four year old culturally disadvantaged children were enrolled in three programs.

1. The major goals of a *traditional* nursery school were to promote the personal, social, motor, and general language development of the children. Teachers capitalized on opportunities for incidental and informal learning, encouraged the children to talk and to ask questions, and stimulated their interest in the world around them and especially in books.
2. A highly structured program (*Direct Verbal*), under the direction of Dr. Carl Bereiter and Siegfried Engelmann, represented a sharp break with the child development tradition. The educational program was derived from an analysis of material to be learned, and it consisted of sessions of intensive direct verbal instruction in language, reading, and arithmetic.
3. A second highly structured program (*Amelioration of Learning Deficits*), under the administration of the research director, Dr. Merle B. Karnes, emphasized content chosen from school-related curricula. Specific tasks were designed to remediate learning deficits by promoting language and cognitive development. An essential feature of the program was to present content in a game format which employed manipulative, multisensory materials, but was structured to require concurrent verbal responses.

The children who participated in these preschool programs have now completed first grade. During this three year period, data were collected on intellectual functioning, language development, perceptual development, school readiness, and achievement. At the end of the first grade, children from the two highly structured programs were functioning above grade level on standardized achievement tests and were significantly higher than the children from the Traditional Group. These findings contrast with those of other preschool research where differences between groups have tended to disappear prior to the completion of first grade.

FIELD-TEST OF THE BEREITER-ENGELMANN PRESCHOOL CURRICULUM IN A SIX-WEEK HEADSTART PROGRAM, *Bruce Rusk*, Ontario Institute for Studies in Education

The purpose of the study was to determine whether in a six week summer Headstart program children following an academically oriented curriculum can make significant cognitive gains over children following a less structured curriculum. Of seventeen Headstart centers, each with approximately fifteen children, eight were designated as experimental and were matched according to socioeconomic level with eight control schools. One control school was not matched. The children were pretested in the first two days of the program and posttested in the last two days on the Pre-School Inventory and Concept Inventory. A *t*-test was applied to the mean difference in gain from pretest to posttest between the matched experimental and control schools. On the Pre-School Inventory the results approach significance at the .10 level. On the Concept Inventory the results were significant beyond the .01 level.

A Pearson product-moment correlation and a *t*-test applied to the pretest results indicated that the matching of the schools was better for the Concept Inventory than for the Pre-School Inventory. It was, therefore, decided to carry out an analysis ignoring the pairing of groups and using the individual child as the unit of sampling. For this purpose a ranking statistic was employed. A Wilcoxen Test for Two Matched Samples applied to the results of the ranking indicated that the results were significant at the .001 level on the Pre-School Inventory and at the .02 level on the Concept Inventory. On both tests a significant number of children in the experimental group went up in rank in relation to the children in the control group.

From these results it was concluded that children in a six week summer Headstart program using an academically oriented curriculum can make significant cognitive gains over children in a program with a less structured curriculum while still receiving the benefit of the health, social, psychological, nutrition, and parent programs and field trips recommended for Headstart.

SUB-CULTURAL VARIATIONS IN PRESCHOOL CURRICULUM NEEDS, *Prof. Roman B. Aquizap,* West Virginia University

Traditional nursery and preschool curricula are inadequate to remedy developmental deficit related to various conditions encountered in the many poverty sub-cultures. The traditional orientation lacks the emphasis necessary for shaping basic cognitive and verbal intellectual behavior absent in many of the homes of poverty areas.

Repeated measurement of the total population of year around and summer Headstart programs in an extremely culturally deficient and geographically isolated community in southern Appalachia is reported. Results of these measurements (using the Cattell Culture Fair Test, Scale I) show no improvement from year to year over a four-year span, with an extraordinarily high number of scores in the defective range; moreover, children did better in some "culture confounded" tests than they did in some of the "culture free" tests. A reanalysis of the results, based upon recomputing Culture Fair and Culture "Foul," I.Q., and Mental Age, showed that these unpredictable results could be accounted for by the special effects of the environment. It also appears that Cattell's "culture free" tests are not necessarily culture free in all sub-cultures.

The results further showed that there were significant differences between the recomputed sub-culture Fair and Foul scores and the original scores. In addition, the poorer performance on the sub-culture Foul scores reflected the local environment deficiencies as determined empirically in a collateral study.

These findings suggest important implications for determining subcultural variations in the preschool curriculum. Some of these necessary changes are discussed, e.g., the shaping of various cognitive skills (as were shown to be deficient on testing) through the use of "free play" incidents as well as by the more formal programmed presentation devices.

Session 10.7 **Division A**

PERSONNEL ADMINISTRATION

Theodore Reller, University of California, Berkeley, Chairman

VARIABLES AFFECTING DECISION-MAKING IN THE SELECTION OF TEACHERS (color film), *Dale L. Bolton*, University of Washington

The purpose of the study was to determine the effects of four information-format variables on the decision process for teacher selection. An experiment was conducted in a simulated situation, and subjects made decisions regarding fictitious applicants for a hypothetical position.

Variables manipulated were: (1) instructions for processing information (either instructions or no instructions), (2) number of documents (multiple or single), (3) masking of information (considerable masking, partial masking, and no masking), and (4) interview information (audiovisual interviews filmed for control purposes, the sound track from the film alone, and none). Measures used to determine the effect of these variables were (1) consistency of decisions (on both an estimate of the behavior of each applicant and a rank ordering of the applicants), (2) fineness of discriminations made, (3) time required to make decisions, and (4) feeling of certainty regarding the estimates of behavior and rank ordering.

The design was a 2x2x3x3 factorial for which analysis of variance was used in analyzing data for each of the four dependent variables. The subjects were 144 elementary school principals selected randomly from the metropolitan area of Seattle.

The results indicated that each of the independent variables affected one or more of the dependent variables, supporting the hypothesis that the format of the information does affect decisions in the selection of teachers. The optimum format consisted of instructions regarding the processing of information, a single summary document, no masking of information, and interviews with audiovisual stimuli; however, audio information (such as acquired in a telephone interview) adversely affected only the feelings of certainty and could, therefore, be recommended where administrators deem it necessary. Whereas the results have direct implications for practice, the materials have significant implications for teaching administrators.

ELEMENTARY TEACHER SELECTION AND EVALUATION POLICIES: A STUDY OF TWO MODELS, *William L. Duff, Jr.*, Washington Internships in Education; *Samuel R. Houston*, Colorado State College

The purpose of the study was to examine the teacher selection policy of a suburban elementary school district and contrast it with teacher evaluation policy.

Two policy models were captured by utilizing *J*udgment *AN*alysis (JAN) techniques. JAN is a multiple linear regression approach used to determine the relative importance of the various profile items for each judge in the expression of his policy.

In the first model, called the *ex ante* model, the hiring policies of four school district personnel charged with teacher selection were captured on the basis of judgments given by each school official. Each official was presented with profile data on 52 teachers. The 32 profile variables consisted of information typically available when teachers are hired, including (1) biographical information, (2) interview data, (3) college preparation, (4) educational experiences, and (5) other experiences.

In the second model, called the *ex post* model, the performance policy of four administrators-teachers was captured via JAN techniques with evaluations of 78 teachers.

Results indicate that, in the *ex ante* model, a policy was clearly stated by each of the judges (approximately 90 percent of the variance was explained). The most important subset of variables was interview data, followed by other experiences, college preparation, educational experiences, and biographical information, in that order. Policy existed in the *ex post* model, though not so clearly defined (approximately 55 percent of the variance was explained). College preparation was the principal source of predictor information in the *ex post* model, followed by other experiences, educational experiences, biographical information, and interview data.

PATTERNS OF PROFESSIONAL GROWTH IN HIGH AND LOW INCENTIVE SCHOOL DISTRICTS, *Clifford P. Hooker*, University of Minnesota

Every major school district in the United States offers a salary increase to teachers willing to earn a Master's degree. If one accepts the contention that graduate study serves a valid professional purpose, and therefore should be encouraged, the question arises whether greater motivation to pursue study is achieved under conditions of relatively higher salary incentive. This question, stated in hypothesis form, suggests that the amount of post-graduate education of elementary and secondary teachers is directly related to the salary policy of the employing district.

This hypothesis was tested using a stratified random sample of "high" and "low" incentive districts in 11 major metropolitan areas. Questionnaires were distributed to a random sample of 6,251 teachers in 31 districts. A total of 5,067, or 81 percent, of the teachers responded. All usable subjects were included in a regression analysis to determine if, holding certain personal and professional characteristics constant, the number of graduate credits earned by teachers can be predicted simply by knowing whether their district of employment is "high" or "low" incentive.

The findings reveal *no* significant difference in the number of credits attained by individual teachers in high and low incentive districts when personal and professional characteristics are controlled. The significant difference is between sexes. Males are more likely to pursue graduate study.

Larger increments between lanes on the salary schedule within the limits represented in the schools under study do not motivate teachers to pursue graduate training. This conclusion lends a measure of discreditability to a widely held popular conception. However, the dollar distance between the B.A. and M.A. lanes was relatively small even in the high incentive districts. No district in the study offered as much reward for the fifth year of study as was paid for each of the first four years of preparation.

THE NON-BEHAVIOR OF SUPERVISORS, *Arnold J. Falusi*, Ontario Institute for Studies in Education; *John C. Croft*, Ontario Institute for Studies in Education

A review of the literature about supervisory behavior disclosed a body of information largely disconnected and hortatory, clearly in need of organization. Information was obtained from an investigation of empirical studies of supervision extracted from eight reference sources and forty-five journals and periodicals covering a span of twenty years. A classification system, inspired and guided by the work of McGrath and Altman (1966) and Foa (1955), was utilized in organizing the findings of ten studies. A procedure was also formulated to determine the least number of studies required to summarize the empirical knowledge about supervision.

Results show that only 10.6% of the information was non-redundant, and that the bulk of knowledge was covered once any twenty studies were reviewed. General conclusions indicate that

1. the classification scheme was useful;
2. there is too much repetition;
3. many findings were largely descriptive and can best be described as normative listings of amorphous role incumbents;
4. most studies examined referred to non-behaviors or duties of supervisors and consequently there was a dearth of specific concepts, thus obviating the possibility of any theoretical integration.

Session 10.8 **Division C**

TEACHER CHARACTERISTICS NO. 1

William W. Lynch, Indiana University, Chairman

ENCODING ABILITY IN TEACHER-STUDENT COMMUNICATION GAMES, *Donna Crossan* and *David R. Olson*, The Ontario Institute for Studies in Education

Attempts to characterize effective teaching have typically ignored the role of a teacher in the communication process. Since transmission of information is central to teaching, it is possible to view the teacher as an *encoder* of a *message* which is subsequently *decoded* by a student. A recent series of experiments using this model have indicated that teachers differ in their ability to encode a standard message. This variability corresponds to differences in the extent to which their message can be decoded by the children. The purpose of this particular experiment was to determine a set of factors which would account for encoding ability, and thus predict effective teaching.

In the first half of the experiment, 35 teachers taught a verbal task and a geometric task to a Grade 6 and then to a Grade 12 student with all orders counterbalanced. Four stooge decoders were used in place of real students, under no feedback conditions: that is, they sat behind a screen and were instructed not to talk to *S*. For the verbal task, *S* taught a list of 12 words for recall in one hour, and for the geometric task, *S* described a series of 4 connected rectangles so that the decoder could draw them. All *S*'s instructions to the decoders were taped. The following information was also obtained for each *S*: academic background, teaching experience, intelligence, and conceptual level.

In the second part, each encoder's verbal, then geometric task instructions were played to a group of 3 students of the appropriate grade and sex—a total of 210 students. After drawing the rectangles, each student wrote down as many of the 12 words as he remembered. Thus, there were two objective measures of the teacher's encoding ability.

It was found that encoding ability as assessed by the performance of the decoders varied with the content of the message and with practice in teaching the task. A linear regression analysis was performed to identify which independent variables accounted for encoding ability.

THE RELATIONSHIP BETWEEN ACHIEVEMENT AND BOYS' AND GIRLS' PERCEPTION OF THEIR RELATIONSHIP WITH MALE AND FEMALE TEACHERS, *Ronald Edwin Hull*, University of New Mexico

The purpose of the research was to test empirically the theory that children at the preadolescent stage of development will evidence greater achievement if they perceive their relationship with their teacher as being more rewarding. More specifically, the question of whether it is more rewarding for boys to be taught by male teachers in the fifth and sixth grades was investigated.

Forty-seven male and forty-eight female sixth grade teachers' classrooms in the Albuquerque Public Schools constituted the sample (2,672 students).

Analysis of covariance was used to determine differences at the .01 level of significance between boys' and girls' achievement in male and female teachers' classrooms. Analysis of variance was used to determine differences at the .01 level of significance between boys' and girls' perception of their relationship with male or female teachers.

It was found that teacher-pupil relationship correlates positively with achievement. Boys and girls both found male teachers to be more psychologically rewarding than female teachers. However, the data showed no significant difference between boys' and girls' achievement under male or female teachers when I.Q. was held constant. Boys achieved significantly better in science than girls and girls achieved significantly better in language arts than boys regardless of teacher sex. The greatest difference lies between boys' perception of female teachers and girls' perception of male teachers, with the former being very unfavorable and the latter being very favorable. The study tends to support the theory that more male role models are needed for students at preadolescent stage of development.

EFFECT OF SELECTED VARIABLES ON TEACHER EXPECTATION OF PUPIL SUCCESS, *James R. Sanders*, Laboratory of Educational Research, University of Colorado; *William L. Goodwin*, Bucknell University

Considerable emphasis is currently being placed on the role that perceptions and expectations play in determining human behavior. The question of what variables are active in establishing a teacher's expectations for a pupil was the focal point of this study.

Seven experiments were designed using each of seven factors found in pupils' cumulative files as one variable and grade level as the other variable, resulting in six 3 x 2 and one 2 x 2 complete factorial designs. The seven factors which were systematically varied, each in a separate experiment, were IQ, course grade average, sex, chronological age, standardized test performance, anecdotal records, and social-economic background. The dependent variable consisted of six questions related to the expectations for the student described by the cumulative file. In addition, a follow-up study was employed to obtain rank orderings of the teachers' weightings of the factors at each of two grade levels and to measure the degree of misunderstanding of the significance of age and IQ factors by teachers. The dependent variable for the follow-up study consisted of instructions to the teachers to rank order the factors, by importance, for each grade level and to answer five questions concerning the significance of various IQ levels and chronological age in determining pupil achievement.

On the six questions concerning cumulative files, significant F-ratios were found for IQ, course grades, standardized test performance, and social-economic background factors. In addition, the grade level factor was found to be a significant source of variance in six of the seven experiments. Teachers perceived social-economic background to be the most important factor in predicting pupil performance and adjustment in the first grade and standardized test results to be the most important factor in predicting performance and adjustment in the sixth grade. The data indicate that teachers perceive IQ, course grades, standardized test performance, and social-economic background as being highly related to the behavior that they should expect from the student. The social-economic background of the pupil appeared to most influence teacher expectations.

Session 10.13 **Division D**

STATISTICAL METHODS

Howard W. Stoker, Florida State University, Chairman

THE APPROXIMATE SAMPLING DISTRIBUTION OF THE STRATIFIED-ALPHA GENERALIZABILITY COEFFICIENT, *Robert L. Mendro* and *Gene V. Glass*, University of Colorado

Feldt (1965) investigated an approximation to the sampling distribution of the Kuder-Richardson 20 reliability coefficient. From this approximation, he was able to derive (with the help of empirical sampling results obtained by Baker, 1963) probability statements about sample values of the coefficient. In the same year, Rajaratnam, Cronbach, and

Gleser presented an extension of their alpha coefficient, called the stratified-alpha coefficient of internal consistency, (α_s), designed to estimate the internal consistency of stratified tests. The purpose of this paper is to present the theoretical formulation of an approximation to the sampling distribution of α_s and to present the results of an empirical sampling study designed to test the approximation. From this approximation, confidence intervals on α_s are constructed.

Using a technique due to Satterthwaite (1946), an approximate distribution is developed for one of the terms in the formula for α_s. Using assumptions about the distribution of mean squares and independent x^2 variables used in a random effects model analysis of variance, the distribution of the entire coefficient, α_s, is approximated by an F-distribution. An expression is developed for the construction of confidence intervals on population values of α_s.

An empirical sampling study designed to examine the accuracy of the approximation was performed. Employing computer simulation techniques, sampling distributions of α_s were generated from hypothetical populations of items. A Type 12 sampling procedure was used, in which random sampling of both items and persons was simulated. The sampling distributions were then examined, and the approximation to the appropriate F-distributions was found to be quite accurate for all sample sizes generated. Limitations of the generalization of the empirical sampling results are discussed.

Thus, a useful set of inferential statistical techniques is made available for stratified internal consistency coefficients.

THE POWER FUNCTIONS OF SOME PROPOSED TESTS FOR THE SIGNIFICANCE OF COEFFICIENT ALPHA,

Anthony J. Nitko, University of Pittsburgh; *Leonard S. Feldt*, University of Iowa

The power functions of two recently proposed significance tests involving Cronbach's coefficient alpha (1951) were investigated. Power function expressions were derived for (1) the one-sample test of the hypothesis $H_\sigma : \rho_\alpha = \rho_\sigma$ (Feldt, 1965; Kristof, 1963), and (2) the two-sample test of the hypothesis $H_\sigma : \rho_{\alpha_1} = \rho_{\alpha_2}$ (Feldt, in press).

The empirical part of this study was concerned with the robustness of the theoretical model developed for the determination of the power of the two proposed tests when the assumptions underlying the model are violated. The results of these robustness experiments led to the following conclusions:

(1) The model proposed for the sampling distribution of coefficient alpha is quite robust, even with dichotomous item scores and with coefficient alpha reduced to KR20, but actual probability of a Type I error is lower than the model suggests. (2) The model predicts the power function of the one-sample test quite closely, but there is a tendency for the model to slightly overestimate the power of the test when the hypothesis and alternative are numerically close and to slightly underestimate the power when the hypothesis and the alternative are numerically distant. (3) The model agrees quite closely with the sampling distribution of the two-sample statistic $W = (1-r_{\alpha_2})/(1-r_{\alpha_1})$, but again the test based on W yields greater control over Type I error than the model suggests.

(4) The model predicts the power of the two-sample test quite well when KR20 is employed, but there is a tendency for it to slightly underestimate the actual power of the two-sample test.

In general, the proposed one-sample and two-sample tests function quite adequately as approximate tests when KR20 is employed, offering good control over Type I and Type II error rates. In particular, the two-sample test is quite robust.

A VARIANCE STABLE RANK METHOD OF SCALING, *Peter Dunn-Rankin*, University of Hawaii

A variance-stable method of scaling paired comparison data is introduced. Tests of significance, determined by the range of rank totals, form an integral part of the scaling technique. Other features of the method include sample size determination and a logical frame of reference. The technique promises to be valuable because of its accuracy, simplicity, and applicability.

A linear transformation of the rank totals that accrue to each object from the *paired-comparison* frequency matrix is proposed as the scale scores for the objects. Tests of significance are then applied to the rank sum differences depending upon the objectives of the scaling process.

Where the Bradley-Terry (1952) and Thurstone (1927) procedures yield scales based on a normalizing transformation (as in taking z-scores corresponding to percentages treated as normal areas), the rank scale is *variance-stable*, as is, e.g., the arcsine transformation of percentages. This is immediately apparent from the fact that any given rank-sum difference has the same probability of occurrence wherever the rank totals may be located on the scale. Thus the variance for all scale values are equal, something that other methods assume but do not guarantee. Moreover, with a given number of judges and items, scale variances are independent of the nature of the items being scaled. Therefore the technique lends itself to scaling a wide variety of stimuli.

Use of multiple comparison tests of significance would appear to be an important adjunct to the scaling process for the following reasons:

1. They provide for making decisions on whether to consider two items as coming from the same population of stimuli.
2. They determine categories of stimuli which can be considered discrete.
3. They can be used to build an index of scalability (SI) for psychological objects.
4. They provide a way of ascertaining the proper sample size for instrument development.

ITEM PARAMETER ESTIMATES FOR THE CASE OF POLYCHOTOMOUS RESPONSE, *Frank B. Baker*, University of Wisconsin

The usual item analysis procedure employed with items allowing multiple response has been to apply successively the techniques for dichotomous response. A given response is treated as the correct response, all other responses pooled as the incorrect response, and the item parameters are estimated. Then the next response is considered the correct response and the process repeated. It has long been recognized that such a practice was incorrect, but a better technique was not available. This gap in item analysis methodology has recently been closed by Baker and Gurland (1968), who derived a maximum likelihood procedure for estimation of the sets of item parameters for items allowing multiple (polychotomous) response. The MLE procedure for the polychotomous case yields simultaneous maximum likelihood estimates of the parameters X_{50} and β for each of the several item characteristic curves involved. It is of considerable interest to compare the item parameter estimates yielded by the older inappropriate procedure with those yielded by the new technique.

Monte Carlo procedures were used to generate item response data from item characteristic curves with specific population values of X_{50} and β. The generated item response data were then analyzed using the two procedures. The sampling mean and variances of the item parameter estimates obtained under each procedure were also obtained. The means were used to study the sampling bias of the two techniques. The sampling variances were used to study the relative efficiency of the two estimation procedures. In addition, the formulas for the asymptotic variances of the item parameters are available for both procedures; hence the ratio of these variances was obtained for a range of item parameter values.

Preliminary results suggest that only small differences exist between the asymptotic variances under the two procedures. Initial Monte Carlo results suggest that when small samples are employed the polychotomous procedure is more efficient.

Session 11.4 **Division C**

SOCIAL STUDIES INSTRUCTION

Jack L. Nelson, Rutgers University, Chairman

ANALYSIS AND EVALUATION QUESTIONS—THEIR EFFECT UPON CRITICAL THINKING, *Francis P. Hunkins*, University of Washington

This study sought to determine whether a dominant use in social studies text-type materials of analysis and evaluation questions, as defined by Bloom's *Taxonomy*, would effectively stimulate the development of sixth grade pupils' critical thinking.

Two hundred and sixty pupils enrolled in eleven sixth grade classes of a large suburban school system served as subjects. Classes were randomly assigned to either experimental Condition A or Condition B. The *Social Studies Inference Test* (Taba, 1964) served as the criterion measure of critical thinking and was administered prior to and immediately following the experiment.

During a four week experimental period, specially prepared materials were used by pupils in each treatment condition. Materials employed in Condition A contained a dominant emphasis (50 per cent) on analysis and evaluation questions. For Condition B, these materials contained predominantly knowledge questions (90 per cent). Each set of materials had a corresponding answer sheet with which pupils checked their answers.

The *Social Studies Inference Test* produced four sub-scores: inference, caution, over-generalization, and discrimination. Analysis of covariance was the principal method utilized in analyzing the data.

Results

Covariance analyses revealed that pupils' use of analysis and evaluation questions did not produce significant differences in critical thinking when compared to pupils' use of knowledge questions. No significant differences were observed between the two conditions for three of the four sub-tests: inference, over-generalization, and discrimination. However, the high-level questions did produce significantly more caution in pupils than did knowledge questions. Reading ability exerted a significant influence on pupils' discrimination scores in both treatments. Reading level significantly interacted with treatment in both caution and over-generalization.

The results of this study indicate that the use of high-level questions (analysis and evaluation) did not stimulate critical thinking in pupils to any greater degree than was true of pupils who experienced knowledge questions.

Several plausible reasons are offered to explain these results.

1. The *Social Studies Inference Test* may not have been an adequate measure of critical thinking.
2. The experiment's duration may have been insufficient to stimulate such thinking.
3. The lack of pupil opportunity to discuss the questions may have tended to inhibit this type of thinking.
4. Critical thinking may be an ability, or abilities, which cannot be taught, but only improved to some unknown degree.
5. Reading which significantly influenced the results may have caused pupils in the two treatment conditions to perform equally as well on the criterion test.

IDENTIFICATION OF "SOCIAL ISSUES" TEACHERS, *Jo A. Sweeney, Byron G. Massialas,* and *Nancy Freitag,* University of Michigan

This paper represents the first phase of an Office of Education research project designed to develop a cognitive category system. This phase of the project consisted of a mailed survey questionnaire identifying teachers of controversial social issues. A probability sample of biology, English, and social studies teachers received the questionnaire. The sample consisted of 682 teachers. Seventy-three percent of the sample returned the questionnaires.

The teachers responded to such items as the issues they considered controversial, time they devoted to such issues, fact and opinion matrix (differentiation between fact and opinion statements), issues they felt should and should not be discussed in the classroom, materials they preferred when teaching social issues, and so on.

Social studies teachers spend the most time discussing social issues, but biology and English teacher responses indicate that social issues are not considered the exclusive domain of the social studies. Race relations, Vietnam, artificial insemination, and birth control are considered highly controversial issues by most teachers. Although race relations and Vietnam are considered highly controversial, the majority of teachers considered them acceptable for classroom discussion, while artificial insemination and birth control are much more "taboo" classroom topics. Although teachers do not spend a large portion of their class time discussing social issues, our data indicate that it is not fear of sanctions which prevents issue discussion. Teachers checked on a list of topics which they would not discuss and indicated reasons why they would not discuss them. Over 80% of the teachers stated willingness to discuss at least 10 of the 13 topics listed; 42% indicated that they would discuss all issues. The perceived sanction-prone (i.e. administrative, community, or parental disapproval) topics are pornography, biological evolution, birth control, and artificial insemination. Teachers who refrain from discussing certain issues stated they do so primarily because of (a) immaturity of class and (b) non-pertinence to subject matter.

Attitudinal items indicate that social studies teachers have a significantly higher belief in student expression. Social studies teachers and male teachers have a significantly higher belief in revealing personal positions during issue discussions. Belief in student and teacher expression is positively related both to time spent discussing issues and to a willingness to discuss all controversial issues.

EFFECT OF VARIETY ON THE LEARNING OF A SOCIAL STUDIES CONCEPT BY PRESCHOOL CHILDREN, *Lynne Schwab* and *Carolyn Stern*, University of California, Los Angeles

This investigation attempted to relate theoretical aspects of the effect of variety of a sample unit in the social studies curriculum. Differential effects of varied vs. repeated training were studied. Head Start children, five years of age, were trained on categories of the concept "work." Eight categories of "work" were presented by means of pictures of people doing the specific kind of work represented by the category.

The following hypotheses were tested:

1. Children trained with intermediate variety—large variety of instances and small variety of categories—will demonstrate superior performance on a transfer to new instances test (Near Transfer).
2. Children trained with low variety—small variety of repeated instances and small variety of categories—will demonstrate superior performance on a mastery test.
3. Children trained with high variety—high variety of categories and small variety of instances—will demonstrate superior performance on a transfer to new categories test (Far Transfer).

Training (2 days) consisted of multiple choice instructional booklets. Testing (3 days) consisted of matching to sample booklets on mastery, transfer to new instances (Near Transfer), and transfer to new categories (Far Transfer).

By analysis of covariance the scores of the intermediate variety treatment were significantly higher ($p < .05$) on the Near Transfer Test, thus supporting Hypothesis 1. No support was found for the two other hypotheses. Children in the three treatments scored approximately the same on the Mastery and Far Transfer Test. It was suggested that an extended instructional program may result in greater treatment differences.

Session 11.5 **Division C**

VERBAL LEARNING NO. 3

Ellen Bouchard, University of Michigan, Chairman

RESPONSE-ASSOCIATED VS. STIMULUS-ASSOCIATED FEEDBACK IN PAIRED-ASSOCIATE LEARNING, *J. M. Knutson* and *J. L. Dunham,* The University of Texas at Austin

Most research in paired-associate learning has focused on conditions in which feedback has provided the correct response to a stimulus. However, in an interpersonal or instructional situation, feedback often takes the form of providing the appropriate stimulus for an incorrect response. The purpose of this study was to investigate the results of providing differential feedback by utilizing the flexibility made possible by a computer-administered paired-associate task.

A series of experiments was conducted to compare the effects of response-associated feedback (RF) and stimulus-associated feedback (SF) in paired-associate learning. In the RF condition, *S* is presented with the stimulus with which his response should be paired, whereas under SF, *S* is given the correct response for the stimulus to which he has just responded. The difference between RF and SF is manifest when an intrusion error is made. Therefore, by manipulating variables which would increase the likelihood of intrusion errors, differences between the two conditions should be enhanced. Consequently, the effects of the two feedback conditions were investigated under different conditions of stimulus similarity, response similarity, and list length.

Using standard paired-associate procedures, all materials were presented on IBM 1510 cathode ray tube terminals connected to an IBM 1500-1800 system. Following *S*'s response, feedback was given immediately according to the experimental condition. In the RF condition, the stimulus paired with *S*'s response was displayed; in the SF condition, *S* was given the correct response to the stimulus tested. There was no feedback difference between these two conditions if *S*'s response was either correct or an extralist error.

The major results of these experiments indicated that correcting an intrusion error by giving *S* the stimulus associated with his response, instead of the correct response, decreased intrusion errors when the likelihood of such errors was high. However, this reduction in intrusion errors did not necessarily result in better learning, but was accompanied by inferior response learning, as manifested by an increase in extralist responses.

EFFECTIVENESS OF REHEARSAL AS A FUNCTION OF AVAILABILITY OF ASSOCIATIVE COMPONENTS, *Ronald E. Johnson,* Purdue University

The effectiveness of rehearsal was hypothesized to be a function of the availability of the associative components of the learning task during rehearsal. After learning a list of paired-associates to a criterion of 7 out of 12 correct, one experimental group had access to a list of the stimuli, a second group had access to a list of the responses, and a third group had access both to the list of stimuli and also to the list of responses. A control group read jokes during the 5 min.-24 sec. rehearsal period.

After the interpolated activities, each of the 4 groups of 22 college *S*s was given five criterion trials on the original list.

The stimuli were two-digit numbers of medium-level meaningfulness and the responses were meaningful three-letter nouns. Five basic 12-item lists were formed, and each list was displayed in four different serial orders. The lists were presented at a 2:2-sec. rate with an intertrial interval of 6 seconds. Standard paired-associate instructions preceded the initial learning period; the rehearsal instructions were not given until the beginning of the rehearsal period.

Analyses of variance provided evidence that the treatment groups were equivalent in learning performance prior to the rehearsal period, $p < .05$. A similar analysis of performance on the first post-rehearsal trial provided evidence that the Stimulus Present and the Stimulus-and-Response Present groups were superior to the Response Present and to the Joke Reading groups, $p < .05$. Analysis of variance of performances on all five post-rehearsal trials, $p < .01$, and on the last four post-rehearsal trials, $p < .05$, supported the view that the effects of rehearsal persisted beyond the first criterion trial. Overall, the results of this experiment provide evidence that the effectiveness of rehearsal is a function of the availability of the associative components. Within the boundary conditions of the present experiment, the availability of stimuli was more critical than the availability of responses.

THE EFFECTS OF STRESS AND CONTEXTUAL CHANGE ON SERIAL LEARNING AND FORGETTING,

Randy Massey, Larry Jensen and *Chris Anderson,* Brigham Young University

This experiment investigated the effects of stress and contextual change on learning and forgetting. Ten nonsense syllables were presented serially on a memory drum. Forty undergraduate females were randomly assigned to one of the following learning-recall sequences: (1) learning and recall under stress, (2) learning and recall under non-stress conditions, (3) learning under stress and recall under non-stress, (4) learning under non-stress and recall under stress.

Stress was induced by instructions relating the learning task to the person's intelligence. In the stress condition *S*s were also told that a tape recording would be used to analyze performance, that attached wiring would provide autonomic and other physiological data, and that the experimenter would observe and record the *S*s' mannerisms. In the non-stress condition the *S*s were told to relax, and that the task was not a test. No equipment was attached to them. Subsequent interviews and a questionnaire showed that a stress and "a non-stress condition" had been created.

For both the acquisition and relearning sessions the number of trials to reach a criterion of six out of ten correct anticipations were recorded. The number of correct anticipations on the first trial of relearning was also analyzed. A two way analysis of variance found no difference between the groups in the number of trials required to reach six correct anticipations during the first session. However, learning under stress resulted in significantly less ($p < .001$) correct anticipations on the first trial of relearning. These same *S*s also required a greater number of trials to reach criterion during relearning ($p < .05$). It was concluded that learning under stress does not facilitate original learning but produces more forgetting in both stress and non-stress recall situations.

AN EXAMINATION OF SOME RELATIONSHIPS AMONG PRODUCTION MEANINGFULNESS, INTELLIGENCE, AND PAIRED ASSOCIATE LEARNING, *Thomas W. Mustico,* State University of New York, State University College at Oswego

This study is based on the assumption that a major determinant of learning success is the amount of previous achievement that is relevant to a new task. In view of this assumption, intelligence test scores tend to predict learning because they reflect a learner's general level of achievement. The more specific a test is to a learning situation, the more accurate will be the predictions that are based on the test scores.

For the purpose of this study, meaningfulness (m) was considered to be a specific achievement that is relevant to the response learning stage of the PA learning task. Hypotheses were generated predicting that an increase in m would occur along with increases in age, IQ, and socioeconomic status (SES). It was also predicted that at extreme conditions of $\overline{m}$ of the learning materials there would be no difference between the mean learning scores of high versus low IQ subjects.

Data were collected from 292 subjects from grades 3, 6, 9, 12, and college sophomores. Trigrams in the form of CCC (three consonants) and CVC (words and non-words) were selected on the basis of determined $\overline{m}$ values (mean number of associations by subjects). Three paired-associate learning lists containing eight S-R pairs were presented "tachistoscopically" to subjects ten times by means of a tape-recorder-activated slide projector. Measures of learning were taken at the end of each presentation of a list.

Correlational analysis revealed significant correlations between meaningfulness and age (.70), SES (—.17), and IQ holding age constant (.38), and a significant decrease in the relationship between IQ and learning as $\overline{m}$ increased. An analysis of variance, followed by an inspection of mean differences lead to the conclusion that at the upper extreme of $\overline{m}$, differences in learning scores for significantly different groups (on the IQ variable) tended to become non-significant. The other extreme (low $\overline{m}$) revealed a somewhat weak tendency for differences to also decrease.

Session 11.6 **Division A**

ADMINISTRATION AND RESEARCH IN HIGHER EDUCATION

John Mallan, The Urban Institute, Chairman

THE WELL-READ TRUSTEE, *Julian F. S. Foster,* California State College, Fullerton

Early in 1968, Educational Testing Services (Princeton, N.J.) conducted a survey of a large sample of college and university trustees. An eight-page questionnaire was constructed and mailed to 10,036 individuals, representing the total membership of the governing boards of 536 institutions of higher education. Of those questioned, 5,180 (52.5%) responded.

One question asked the respondents to indicate their familiarity with each of fifteen major works on higher education. Kerr, *The Uses of* *the University,* Goodman, *The Community of Scholars,* Millett, *The Academic Community,* and Perkins, *The University in Transition* were among the works included. Five choices of response were offered for each title: "Have read completely," "Have read portions," "Have briefly examined, but not read," "Know of the book, but haven't seen it," and "Have never heard of it." Numeric values will be assigned to each of these responses (4, 3, 2, 1 or 0) and a total score will be computed for each individual respondent on the fifteen items. Respondents will then be divided into three groups on the basis of this score: high scorers ("well-read"), low scorers ("unread"), and scorers in the middle range, who will be excluded from the research. Preliminary work suggests that the amount of reading reported by trustees is extraordinarily light, with about 25% indicating no familiarity whatsoever with any title.

Another section of the questionnaire required respondents to indicate agreement or disagreement with 29 propositions concerning higher education. Topics dealt with included student protest, academic freedom, teaching effectiveness, patterns of institutional governance, and the role of federal support. Responses of the two groups, well-read and unread, to each of these items will be compared. The results will form a profile of the beliefs and attitudes of each group.

This research should have practical implications for presidents who seek to 'educate' their boards, and for others who wish to assess the role which trustees can best play in higher education.

SOME OBSERVATIONS ON THE RELATIONSHIPS BETWEEN RESEARCH PRODUCTIVITY AND STUDENT EVALUATIONS OF COURSES AND TEACHING, *William M. Stallings,* University of Illinois (Champaign-Urbana); *Sushila Singhal,* University of Illinois (Champaign-Urbana)

Both students and faculty have charged that the reward structure in institutions of higher learning favors the "researcher" over the "teacher." College and university administrators have rejoined that, in proper proportions, teaching and research are mutually fructifying. Since neither side has presented much hard data, this paper offers some tentative observations about the relationships between student evaluations of courses and instructors with indices of research productivity.

Student evaluations of courses and instructors were obtained from the Course Evaluation Questionnaire (CEQ) developed by R. E. Spencer at the University of Illinois. Six subscores (including an Instructor Score) and a Total Score are derivable from this instrument. According to Spencer, the Total Score is the most reliable and most important score. The CEQ is normally administered on a voluntary basis at the end of each semester.

The University of Illinois compiles an annual list of faculty publications. From that document, weighted bibliographic counts were computed for 128 instructors on whom CEQ data were available. These bibliographic counts were used to generate the Research Productivity Index (RPI). An individual RPI score was defined as the number of publications multiplied by the appropriate weights for each category of publication. Two arbitrary weighting systems, designated Key I and Key II, were used.

A correlational analysis was made using the variables of CEQ Total

Score, CEQ Instructor Score, instructor's rank, and RPI. With both keys, academic rank had significant correlations with RPI and CEQ Total Score. That is, higher ranks were associated with greater research productivity and with higher student course evaluations. Only with Key I was there a significant correlation of RPI with CEQ Total Score.

The results of this study lend little support to the assertion that a productive researcher tends to be a good teacher—and vice versa.

EDUCATIONAL ATTITUDES OF COLLEGE AND UNIVERSITY TRUSTEES, *Rodney T. Hartnett*, Educational Testing Service

At a time when much controversy exists over the proper role of college and university governing boards, it is remarkable that so little is known about who trustees are, what they do in their roles as trustees, and how they feel about some of the more pressing current issues in American higher education. The purpose of this research was to gather information which would fill this void, with the hope that such information would provide a much better understanding of the current goals and practices of higher education and supply new insights into the current controversy regarding the governance question.

An eight-page, 200-item, machine-scorable instrument was developed for this survey. After obtaining the names and addresses of trustees from 536 college and university presidents (with institutions selected appropriately according to type of control and level of course offerings), questionnaire responses were obtained from over 5,000 governing board members. All questionnaires were completed anonymously.

The data were analyzed separately by type of control (public, private) and level of course offerings (junior colleges through universities). Perhaps the single most important finding was that trustees, like the institutions they govern, are characterized by great diversity. Those who govern private universities, for example, differ markedly in many ways from those who serve on boards of 4-year state colleges. (Even the simple matter of how much time trustees spend on board-related activities varies widely across different types of institutions.) Other findings are discussed within broad areas of attitudes, such as academic freedom (trustees appear somewhat reluctant to accept a liberal notion of academic freedom), the business-like orientation of trustee (business executives on governing boards do *not* favor "running a college like a business"), the decision-making process, and the trustees' political preference and ideology.

ROLE PERCEPTIONS OF EDUCATIONAL RESEARCHERS AND ADMINISTRATORS RELATIVE TO THE IMPLEMENTATION OF RESEARCH FINDINGS, *Sandra Damico*, The Ohio Education Association; *Desmond L. Cook*, The Ohio State University

The principal objective of this study was to examine the perceptions of responsibility for the implementation of research findings held by educational researchers and administrators. A secondary objective was to compare similarities between patterns of response in the educational community to those uncovered in a comparable study of researchers and managers in the business community.

Four mutually exclusive statements representing different concepts of degree of responsibility for the implementation of research findings were submitted to geographically stratified random samples of approximately 300 AERA members and 300 superintendents of districts having over 6,000 student enrollment.

The results of the study showed that educational researchers and administrators were in general agreement over their respective responsibilities. This agreement focused upon the statement reflecting a need for mutual understanding of each other's role. Responses within the educational community were found to be more in agreement on the need for mutual understanding than was the case for the business community study.

Session 11.7 **Division C**

MATHEMATICS INSTRUCTION NO. 1

Thomas Romberg, University of Wisconsin, Chairman

THE RELATIVE EFFECTS OF TWO FORMS OF SPIRAL CURRICULUM ORGANIZATION AND TWO MODES OF PRESENTATION ON MATHEMATICAL LEARNING, *Jenny R. Armstrong*, The University of Wisconsin

The purpose of this study was to determine the relative effects of two forms of spiral organization, area or topical, and two instructional modes of presentation, inductive or deductive, on the mathematical learning of sixth graders assessed at each of six cognitive levels, within three areas, and on four topics. The mathematical areas chosen for study were set theory, number theory, and geometry. The topics selected for examination were terminology, relations, operations, and properties.

A sample of 228 sixth graders were randomly selected from three elementary schools in the Madison, Wisconsin, Public School System. The sample was composed of pupils from above average socioeconomic levels with average I.Q. of 117. The selected subjects were randomly assigned to each of the four treatment conditions, area-inductive, area-deductive, topical-inductive, topical-deductive. Four teachers were trained to implement each of the four treatment conditions. The teachers were rotated among treatment groups so that each teacher taught each group for thirteen sessions. The experimental program lasted for fifty-two sessions, each forty-five minutes in length, over a three month time period. A 2X2 completely crossed and randomized factorial design was utilized with fifty-seven replicates in each cell. Both multivariate and univariate analyses of covariance were used to analyze the data. All F ratios were evaluated at the .05 level of significance.

The results indicated that the form of spiral organization and mode of presentation differentially facilitated various types of learning when adjustments were made for the pupils' mathematical aptitude, prior (before instruction) mathematical learning, and global intelligence. The area-spiral form of curriculum organization better facilitated mathematical learning at the knowledge level, whereas the topical-spiral form produced better learning at the evaluation level. The area-spiral form of organization resulted in greater number theory learning. The inductive mode of presentation fostered the learning of operations, while the deductive mode resulted in greater learning of mathematical properties. The interaction of curriculum and instruction variables were not found to significantly affect mathematical learning.

A STUDY OF VARIABLES WITHIN THE METHOD OF INDIVIDUALLY GUIDED DISCOVERY IN SECONDARY SCHOOL MATHEMATICS,

Dr. Chester D. Carlow, Ontario Institute for Studies in Education

The purpose of the experiment was to investigate the relationships of the variables of conceptual structure, consolidation after discovery, and learner personality to learning, retention, and transfer in mathematics *within* a clearly defined teaching method of individually guided discovery.

Thirty-six "successful" ninth grade college preparatory males were selected randomly from three schools, and were randomly assigned to two conceptual treatments (permutations-combinations *or* ordered partitions). Three teachers were employed within each treatment, and there were three consolidation levels (supervised practice amounting to 0%, 25%, 50% of the time required to complete the learning task). The learning task consisted of 15 generalizations in probability. After posing questions, the teachers provided hints, with flexibility, to the learners until the student was able to formulate the generalization correctly. Attempts were made to control deviation from the method of guided discovery as defined, task difficulty, the Zeigarnik effect, the Hawthorne effect, obliteration of meaningful learning, superfluous consolidation during discovery, amount of information provided per hint, sense-modality bias, and communication among learners during the study. Retention (three weeks) and transfer were measured.

Consolidation after discovery promoted retention. Intelligence and conceptual level (Hunt-Halverson, *Sentence Completion Test*) were positively correlated with learning and retention. Submissiveness and closeness (Sterns, *Activities Index*) were negatively correlated with performance. The conceptual structures tested, teachers, and the remaining ten factors of Sterns' *Activities Index* were generally not significantly correlated with performance, though some trends were noted.

Careful definition and research on variables *within* the method of guided discovery teaching is prerequisite to good comparison studies involving guided discovery and other methods. The present study reveals the importance of practice and personality, as well as intelligence, for difficult tasks and genuine discovery.

A COMPARISON OF THREE METHODS OF TEACHING SELECTED MATHEMATICAL CONTENT IN EIGHTH AND NINTH GRADE GENERAL MATHEMATICS COURSES,

William D. McKillip, James F. Strickland, Jr., and *Freddy J. Maynard,* University of Georgia

The problem in a pair of studies was to compare three different methods of teaching selected mathematical topics in grades eight and nine. The three methods selected were: (D) non-verbalized student discovery, (E) student-teacher cooperative development, and (S) authoritative teacher-presented statements and examples.

The basic materials selected were the *Experiences in Mathematical Discovery* units, pamphlets published by NCTM, which were re-written in three different forms corresponding to the teaching methods D, E, and S. Teachers cooperating in the study were trained in the methods they were to use, were supplied with sets of materials conforming to that method, and were observed at various times to determine how closely they were following the assigned method. Eighteen eighth and ninth grade classes containing more than 500 students participated in the studies.

A groups-within-treatments analysis of variance design was used to evaluate the three methods of instruction. Some of the experimental results were found to be significant at the .05 level.

A COMPARATIVE STUDY OF THE EFFICACY OF A SELECTED ARITHMETIC DRILL PROGRAM WITH SECOND AND FOURTH GRADERS, *Leonard Pikaart* and *Sue Ellis Nutting*, University of Georgia

This study was designed to evaluate the effectiveness of a slide projector mediated drill program in improving the computational abilities of students in the second and fourth grades. Measures of speed and accuracy for each of six tests of basic facts in addition and multiplication were obtained in pretests and posttests.

The 178 subjects selected for the study were second and fourth grade students in a Georgia school classified as educationally disadvantaged on the criteria in Title I of the Elementary and Secondary Education Act. The experimental group received the drill program in addition to regular arithmetic instruction, although the total time devoted to arithmetic instruction was not increased. The drill program was used over a period of 16 weeks, averaging about one hour of use per week—usually three sessions of approximately 20 minutes each.

Data were analyzed via a Lindquist Type I (Repeated Measures) model. In general, the experimental group improved significantly more than the control group. The experimental group in the second grade exhibited significant improvement in speed for both multiplication and addition.

Session 11.8 **Division D**

STATISTICAL METHODS: ANALYSIS OF VARIANCE AND COVARIANCE

Thomas M. Goolsby, Jr., University of Georgia, Chairman

CORRECTION FOR ATTENUATION: A GENERALIZABILITY APPROACH USING COMPONENTS OF COVARIANCE, *Kenneth J. Travers*, University of Illinois

Correction for attenuation, as traditionally approached, is legitimate within the assumptions of the classical model. But in many investigations these assumptions are not met. Applying the classical correction formula in these cases yields correlations which are open to misinterpretation and may even be theoretically impossible.

The present study approaches correction for attenuation by using the basic concepts of generalizability theory. The emergence of generalizability theory during the past decade, fostered primarily by the work of Lindquist (1953), Cronbach, *et al* (1963) and Gleser, *et al* (1965), has opened up new aspects of the theory of measurement. But whereas this work to date has been univariate, the present study extends the theory to the multivariate case.

Since classical theory recognizes, essentially, only a one-facet experimental design (that of equivalent forms of tests), only one corrected correlation is possible regardless of the data gathering procedure employed. Generalizability theory, however, makes available many different "corrected" coefficients (correlations between universe scores) depending upon the design of the experiment.

The paper gives formulas for the various alternative "corrected" correlations and provides examples from educational settings for basic experimental designs. Similar principles and procedures may be utilized in extending the theory to more complex designs.

AN ANALYSIS OF VARIANCE FRAMEWORK FOR MATRIX SAMPLING: IMPLICATIONS FOR BINARY AND NON-BINARY ITEMS, *Ken Sirotnik*, University of California, Los Angeles

In order to present matrix sampling in a rigorous and generalized framework, Hooke developed an algebra involving symmetric polynomials of the elements in a matrix. Certain linear combinations of the polynomials (bipolykays) turn out to be estimates of the moments of the matrix population, thus providing Lord with a convenient way to obtain formulas for the estimated examinee score mean and variance from a matrix sample.

The purpose of the present paper is to derive these same formulas using a simpler, more intuitive framework. Specifically, this paper attempts to demonstrate the following points:

1. Matrix sampling can be viewed as a simple two factor, random model analysis of variance design, the matrix sampling formulas for estimating the mean and variance being simply the analysis of variance formulas for estimating components of the underlying linear and additive model.
2. In contrast to non-binary items, binary items can be expected to violate homogeneity of variance and independence assumptions; when the between-subjects variance is less than the residual variance, a negative matrix sampling variance estimate results.
3. Given the use of binary data, the usual achievement test items, in contrast to objective-meeting or criterion-oriented test items, may be more amenable to matrix sampling methodology with respect to variance estimation. (Estimation of the mean appears to be unaffected by the type of items or test used.)
4. Of the various attempts to deal with negative variance component estimates in multiple matrix sampling, either the maximum likelihood method of equating these estimates to zero or the method of Winsorizing the distribution of estimates appear to be most promising. A simulation study is necessary to determine the shape of the distribution of variance component estimates for matrix sampling as well as the relative efficiency of the two methods for handling negative estimates.

THE SIMULTANEOUS DETERMINATION OF SAMPLE SIZE AND TYPE I ERROR-RATE IN ANOVA DESIGNS GIVEN THE POWER AGAINST EACH OF TWO ARBITRARY NON-NULLITIES, *Leonard P. Kroeker* and *G. William Walster*, University of Wisconsin

The meaning of statistical significance is at best unclear when the investigator has not chosen intelligently the value of the sample size and the type I error-rate. This paper presents a mechanism by which rational choices may be readily made.

Most researchers realize that it is desirable to have a large amount of power against "important" non-nullities. However, not all of them show comparable concern about the effect of a large amount of power against "trivial" non-nullities. Considerations of too much power against small non-nullities and too little power against "important" non-nullities motivated this investigation of the simultaneous choice of sample size and type I error-rate.

A new measure of non-nullity (Δ) is introduced. This measure is a generalization of a mean difference relative to underlying variability. Unlike traditional non-centrality parameters, Δ can be readily used by the researcher.

The central feature of the paper concerns the determination of a unique power curve for one's statistical test once the minimum amount of power desired against a non-nullity large enough to be "important" and the maximum amount of power desired against a non-nullity small enough to be "unimportant" are specified.

A computer program was written that solves simultaneously for values of the sample size and the type I error-rate associated with the desired power curve. The program is written in FORTRAN IV for the CDC 3600.

A CLASS OF COVARIANCE STRUCTURE MODELS, *William J. Bramble, William H. Schmidt,* and *David E. Wiley,* University of Chicago

A class of sixteen models is presented for analyzing the structural characteristics of covariance matrices. A method of analysis is proposed for these models which operates on the sample covariance matrix. The results include maximum likelihood estimates of the parameters, likelihood-ratio tests of the goodness-of-fit of the model, and large-sample variances and covariances of the estimates.

A general model for the analysis of covariance structures includes three parameter matrices: a matrix of weights for the observed variables in term of the latent variables, the covariance matrix of the latent variables, and a matrix of unique variances. By placing various restrictions on these matrix parameters a class of sixteen models results. The following cases are considered: (1) the matrix of weights may be general (unspecified), general having certain elements specified, completely specified, or completely specified but scaled by an unknown (but estimable) matrix; (2) the covariance matrix of the latent variables may be diagonal or general, depending on whether the latent variables are assumed orthogonal or oblique; and (3) the matrix of unique variances may be general diagonal or diagonal with equal elements, depending on whether the unique variances are assumed heterogeneous or homogeneous.

The less restrictive of the above models are closely related to the traditional factor analysis model. However, the techniques presented here for the solution of these models employ maximum likelihood estimation of the parameters. The most highly specified of these covariance structure models is the mixed model for the analysis of variance. Several of these models have been analyzed by maximum likelihood procedures before. (See Joreskog, 1967; Bock and Bargmann, 1966). We follow Joreskog in using the method of Fletcher and Powell (1963) for the maximum likelihood estimation.

In addition to the new models proposed here, the large structured class allows one to test sequences of naturally ordered models.

A general computer program for the maximum likelihood estimation of the parameters in these models is described. The data from the Miller and Lutz (1966) study of teaching practices and learning situations are reanalyzed according to the procedures.

Session 12.4 **Division C**

INTERACTION ANALYSIS NO. 1

Thomas Samph, Syracuse University, Chairman

STUDENT ACHIEVEMENT LEVEL AND DIFFERENTIAL OPPORTUNITY FOR CLASSROOM PARTICIPATION, *Thomas L. Good*, University of Texas

It was hypothesized that teachers expect different performance levels, and based upon this expectancy, extend different types and frequencies of response opportunities and provide differential feedback to students as a function of achievement level. Subjects were selected from four first-grade classrooms in two predominantly white, working-class neighborhoods.

Target teachers were requested to rank students in order of achievement on the basis of teacher observation of student behavior and performance throughout the year. The first four and last four students on the teacher achievement rankings were respectively classified as High and Low observation groups. The Middle four students were identified by a median split analysis of pupil achievement rankings. The researcher used the Good Opportunity Observation Device and systematically recorded response opportunities and feedback instances provided by the teacher to target students. In total, ten hours of instructional proceedings were viewed in each class.

It was found that teachers consistently and significantly afford Highs more response opportunity and positive feedback information than Lows. It was also found that Lows receive significantly more negative feedback than Highs. The data amply demonstrate that Lows, psychologically, live in a different room than do Highs.

It is suggested that such a response deprivation separates Lows from classroom life and mitigates against their educational progress. It is thus contended that teacher behavior may be a major factor contributing to the phenomenon of cumulative deficit wherein students of low achievement progressively decline relative to their classmates.

THE RELATIONSHIP OF SELF-SUPERVISION TO CHANGE IN SELECTED ATTITUDES AND BEHAVIOR OF SECONDARY STUDENT TEACHERS, *Donald P. Johnston*, Memphis State University

This study compared instances of self-supervision by student teachers, in which they applied the Flanders system of interaction analysis to their own teaching behavior recorded on video-tape with no supervisor present, with instances of supervision of student teachers in a more traditional manner. A stratified random sample of 84 student teachers in secondary school academic subject areas was distributed into four treatment groups according to a trichotomization of their scores on the *Minnesota Teacher Attitude Inventory*. Each subject taught two twenty-minute lessons, with the same content, to different classes of five pupils. Three groups of subjects supervised themselves; the other group was supervised in a more traditional manner (a supervisor was present).

Data were analyzed for relationships between MTAI scores and interaction behavior (determined by the Flanders system) on the first lesson taught, for changes in the incidence of indirect behavior in both lessons by student teachers, for the relationship between observed amounts of indirect control and student teacher estimates of such amounts, and for changes in MTAI scores.

The following conclusions were drawn: (1) no significant relationship exists between attitudes and teaching behavior before supervisory treatment; (2) supervisory treatment tends to promote a significant relationship between attitudes and teaching behavior; (3) self-supervision tends to promote indirect teaching; (4) self-supervision tends to promote higher scores on the MTAI; (5) estimates by student teachers of the percentage of indirect teaching they exhibit in their lessons are very inaccurate under both traditional supervision and self-supervision; (6) no significant relationship exists between time and attitude change in student teachers supervised in a traditional manner.

The method of self-supervision studied would seem to provide a desirable alternative in the supervision of secondary student teachers where indirect teaching and pupil-accepting attitudes are sought.

THE RELATIVE INFLUENCE OF THE GROUP OF PUPILS AND OF THE TEACHER AS DETERMINANTS OF CLASSROOM INTERACTION, *Lester Oppenlander*, Indiana University at Indianapolis

The purpose of this study was to determine the relative influence of the teacher and of the group of pupils in determining the pattern of interaction which transpires in the classroom.

Using Flanders' observation categories, four subject-matter teachers (arithmetic, science, language arts, and social studies) in a departmentalized sixth grade were observed as they interacted with two sections of pupils. All four teachers taught both sections. These two sections were the highest and lowest of five sections grouped according to achievement by the school administration (based on Iowa Test of Basic Skills and Stanford Achievement Test). Each teacher was observed with each section throughout ten class periods. Half of the observations (40) were made very early in the semester; the other half were made late in the semester. Analysis of variance (2x2x4) of sections by time by teachers with five replications in each cell was run using coded i/d ratios as cell entries. Evidence of direction of causation was sought in the interaction effects involving the time dimension.

Both the teacher main effects and the section by time interaction were significant at .05. No other effects were significant at that level. These two groups of pupils exerted more influence on the classroom interaction than did the basic styles of the four teachers. It would seem that these teachers, regardless of their basic styles or preferred way of teaching, were pressured to a classroom interaction pattern by the pupils in the classroom.

The study cannot be used to make any sweeping generalizations, but suggests the need for a larger study with more adequate sampling for more valid generalizations. If these results hold in later studies, emphasis in future studies of classroom ecology should likely shift to characteristics of pupil groups rather than teacher characteristics.

A SOCIAL-PSYCHOLOGICAL STUDY OF CLASSROOM VERBAL BEHAVIOR, *Philip Piele*, University of Oregon

The purpose of this study was to investigate the relationship of teacher open and closed mindedness to classroom verbal behavior.

Seventy teachers from elementary and junior high schools in eastern Oregon, eastern Washington, and Idaho were administered the Rokeach Dogmatism Scale (Form E). Seventeen high scorers and 17 low scorers were selected as subjects for this study. All 34 subjects were in the upper or lower quartiles of dogmatism scale scores. None of these subjects had any previous training in interaction analysis.

Each subject tape recorded six 20-minute segments of regular class lessons. Observers trained in the use of the Flanders' system of interaction analysis listened to the tape recorded lessons and recorded the verbal behavior on tally sheets. A computer matrix plotting program was used to plot and compute the appropriate column and cell totals and percentages.

The significant findings of this study were that closed minded teachers appear to use a greater variety of verbal behaviors and to monopolize classroom talk more than do open minded teachers, and that students of open minded teachers appear to use more extended responsive talk and to verbally interact with each other more than do students of closed minded teachers.

It is suggested that, because closed minded teachers are more concerned about classroom control than open minded teachers, they tend to discourage student talk through the use of a wide variety of verbal behavior. It is further suggested that some of the verbal behavior used by closed minded teachers to control student talk are recorded as indirect influence by the Flanders system. An assumed generic link between Flanders' indirect-direct concept and Anderson's integrative-dominative, Lippitt's democratic-autocratic, and Cogan's inclusive-preclusive is questioned.

Session 12.5 **Division C**

ACQUISITION OF READING SKILLS NO. 3

Walter R. Hill, State University of New York at Buffalo, Chairman

WHEN A WORD ENDS IN VOWEL-CONSONANT-E . . . , *Lou E. Burmeister*, University of Texas at El Paso

Most phonic systems teach, "When a word ends in vowel-consonant-e, the *e* is silent and the vowel is long." Several recent utility level studies have indicated that this generalization has very limited usefulness. The present study was designed to look at a large stock of English words to see if another, more valid, generalization could be formulated to describe the sound, or sounds, of vowels in this situation.

All of the words that ended in single vowel-single consonant-*e* in the 17,310 words of the Hanna, *et al*, *Phoneme-Grapheme Correspondences as Cues to Spelling Improvement* were used in the study. In order to describe the phonic situation for these words, the findings of Hanna, *et al*, were reversed to grapheme-phoneme correspondences.

The phonemes that each vowel represents in this graphic situation are described in terms of both frequency and percent of occurrence of each phoneme for each grapheme so that logical comparisons can be made both within and among the vowels.

Examination of the data indicates that the generalization examined is more valid for some vowels than it is for others. For *a*, *e*, *i*, *o*, *u*, and *y*, it reaches the utility levels of 78.9, 67.4, 61.1, 85.6, 78.3, and 100 per cent, respectively (y = ī).

There are three specific groups of words which form the major exceptions to the generalization. These groups are (1) a-e = ĭ, especially in words ending in -ace, -age, -ate, (2) e-e = â, including *there*, *where*, and *ere*, and compounds using these roots, (3) i-e = ĭ, especially in -tive and -ive words.

There are 67 primary level words which are exceptions.

PRONUNCIATION OF SYNTHETIC WORDS WITH PREDICTABLE AND UNPREDICTABLE LETTER-SOUND CORRESPONDENCES,

Robin S. Chapman, *Robert C. Calfee*, and *Richard L. Venezky*, University of Wisconsin

A child learns to read by first learning to translate from visual symbols to sound. One part of this process is the formation of correspondences between letter patterns and the sounds for which they stand. The major concern of the study reported here was to find the extent to which readers used such correspondences in pronouncing synthetic words, and how they pronounced synthetic words for which no such regular correspondences existed.

A list of synthetic words was prepared for testing the pronunciation of predictable patterns (final *-e*, *c* before *e* and *i*, and *c* before *a*, *o*, and *u*), unpredictable patterns (vowel digraph spellings such as *ai*, *ou*), and miscellaneous spellings (*th*, final *s*). For the predictable patterns, the appropriate responses were taken to be the long form of the vowel for final *-e* patterns (*rate*, *mete*, *bite*, *rote*, *cute*), *c* pronounced /s/ before *e* and *i* (*cell*, *city*, and *c* pronounced /k/ before *a*, *o*, and *u* (*cake*, *coke*, *cute*). A 40-item list and 5 pre-training items, prepared on slides in capital letters, were presented to *S*s in one of two random orders. Responses were recorded and transcribed by graduate students trained in phonetics. Participating in the study were third and sixth graders from two elementary schools, eleventh graders, and college students.

It was found that, to the extent that a child in third grade was identified as a good reader, he showed some mastery of predictable letter-sound correspondences. This mastery increased through high school, but the correlation with reading achievement decreased, presumably because this ability is only one of many necessary for skilled reading. Poor readers made more and "wilder" errors on predictable patterns and gave less consistent responses to unpredictable spellings.

EFFECTS OF INSTRUCTION IN LETTER KNOWLEDGE ON ACHIEVEMENT IN READING IN THE FIRST GRADE, *Dorothy C. Ohnmacht*, University of Missouri

This study investigated the relative effects of the selected ordering of instruction in letter names, letter names and sounds, and sight words on achievement in word knowledge, word discrimination, and comprehension. 208 1st grade Ss were assigned to one of three treatment conditions. Treatment Group A received instruction in letter names followed by the teaching of sight words. Treatment Group B received instruction in letter names and sounds followed by the teaching of sight words. Treatment Group C received instruction in sight words followed by the teaching of letter names and sounds. Comparisons were made for Ss in three treatment groups in the total sample, for sub-samples of Ss on several intelligence levels, and for sub-samples of Ss on several reading readiness levels.

Criteria consisted of tests of word knowledge, word discrimination, and comprehension of sentences and paragraphs. Tests of prior knowledge of letter names, prior knowledge of letter sounds, information and oral language comprehension, and intelligence provided data utilized as covariates. Scheffe's Test was employed to compare adjusted means on criterion scores following rejection of any null hypothesis.

Initial instruction in letter names and sounds produced greater achievement in word knowledge and word discrimination than did initial instruction in sight words. Comparisons at several intelligence levels indicated this general superiority to be a function of significant differences for children of average intelligence only. When comprehension was the criterion there was no general treatment effect detected. However, initial instruction in letter names and sounds produced superior results in comprehension for children who were average on measures of reading readiness. Initial instruction in letter names appeared of value to reading achievement mainly for children below average on measures of reading readiness when word perception was the criterion.

AN EXPERIMENTAL STUDY DESIGNED TO TEST THE RELATIVE EFFECTIVENESS OF A MULTI-MEDIA INSTRUCTIONAL SYSTEM, *Donald R. Senter*, Educational Developmental Laboratories, Inc.

The purpose of this study was to test the relative effectiveness of two methods of teaching reading to more than 1300 students in 56 classroom units at the first year level. The control groups used basal reading programs and the experimental groups used the *Listen Look Learn* multimedia system.

The primary analysis was done utilizing a 2x3x2 (treatment by ability level by sex) factorial design. Analysis of Variance and Analysis of Covariance were performed, and results consistently favored the subjects from the LLL groups for four subtests of the *Stanford Achievement Test*, Primary I, and the *Cooperative Primary Listening Test* at significant levels. The interaction of treatment by ability level was consistently significant and resulted from high achievement scores attained by average range (88-112) children.

A secondary analysis was performed utilizing a 2x4x2 (treatment by community size by sex) factorial design in order to test the relative effectiveness of the two teaching methods in communities of varying size. Subjects from small suburban and rural areas were consistently favored at a significant level over subjects from larger towns and cities, although subjects from LLL groups retained higher achievement levels for all categories of community size.

A second auxiliary analysis was performed following non-stratified blocking on socioeconomic status in order to test the relative effectiveness of the two instructional methods for children of differing backgrounds. Consistent significant differences were detected favoring the high socioeconomic category, although consistent differences between LLL and basal reading groups were found favoring children from LLL groups.

The results of this study suggest that a multi-media approach to reading instruction at the first year level may result in higher achievement in reading and in listening at the completion of one school year than does a basal reading program.

Session 12.7 **Division C**

FEEDBACK AND REINFORCEMENT NO. 1

Willavene Wolf, Ohio State University, Chairman

EFFECTS OF VARYING QUALITY, AMOUNT, AND DELAY OF REWARD IN THE CLASSROOM, *Katherine A. Bemis* and *Glenn B. Schroeder,* Southwestern Cooperative Educational Laboratory, Albuquerque

The purposes of this research were to study and develop classroom management techniques, to increase pupil interest in the learning process, and to contribute to pupil gain in desirable behaviors. A first grade curriculum served as a basis for determining the effects of various teaching strategies and classroom management techniques. Specifically, the parameters of interest were quality of reward and delay of reward.

Two types of reinforcing conditions were compared: the conventional verbal praise of the teacher, and a point system where accumulation of a specified number of points, indicated by a progress chart, results in a tangible academic reward. A short-term incentive goal condition of one week and a long-term goal condition of six weeks were studied. The rewards for the short term were of lesser magnitude than for the long term. Teachers were assigned randomly to experimental and control conditions during each of three six week phases.

Ten maturational classes were compared to the eight experimental classes on affective and cognitive measures. The CAT, Form W, was used to assess cognitive differences. The eight experimental classrooms achieved significantly better ($p < .01$) on Arithmetic Fundamentals subtests than did the maturational controls. The SWCEL Student Questionnaire was employed to assess pupil affect. The experimental groups were significantly higher ($p < .01$) than control groups in the area of self-esteem. Within-group differences indicated that the classes rewarded under short-control-long and long-short-control conditions achieved significantly better ($p < .01$) than those of other reward conditions on affective and cognitive measures.

REWARD PREFERENCE PROFILES OF ELEMENTARY SCHOOL CHILDREN, *G. Phillip Cartwright* and *Carol A. Cartwright*, The Pennsylvania State University

The purpose of this investigation was to determine the reward preference patterns of groups of elementary school children in grades four, five, and six. Information concerning children's preferences for various classes of rewards will be of value to both researchers and classroom teachers in the development of efficient and effective motivation systems for groups of children and for individual children.

The Dunn-Rankin Reward-Preference Inventory was administered to approximately 500 children, both males and females, in grades four, five, and six. The total group was divided into sub-groups according to grade level and sex. The Reward-Preference Inventory is based on the paired-comparisons method and includes sample rewards from five general classes of rewards: (1) adult approval, (2) competition, (3) consumable, (4) peer approval, and (5) independence. Scoring the inventory yields a rank order of preference for the five reward classes for each subject. Scale scores are computed from group rank totals, and group preference profiles are derived from the scale scores.

Reward preference profiles were obtained for each sub-group (grade level and sex). Results of the study indicate adult approval was the most preferred class of rewards for all sub-groups. In general, independence and consumable rewards were least preferred. Extensive individual differences within groups were noted.

FEEDBACK DELAY AND RETENTION IN THE CLASSROOM, *Arthur J. More*, University of British Columbia

The purpose of the study was to investigate the relation between immediacy of feedback and retention in a school setting.

Recent studies by Brackbill, and by English and Kinzer, resulted in greater retention by subjects for whom feedback had been delayed than by subjects who received immediate feedback. However, these studies involved learning materials which were relatively meaningless or subjects who were college students. Furthermore, recent studies suggested a relationship between the effect of feedback delay and the type of criterion.

Subjects, 560 grade eight students, first read one of two different articles, and were tested on the content by a multiple-choice test. Second, subjects were provided with knowledge of the correctness of their test responses after no delay, two and one half hours, one day, or four days delay. Third, subjects were retested with a resequenced form of the first test. One half of the subjects were retested immediately after feedback (acquisition criterion) and the remainder were retested four days after feedback (retention criterion).

A factorial analysis of covariance design was utilized with retest scores as the dependent variable and IQ and reading grade level as covariates.

All main and interaction effects were statistically significant. Acquisition criterion groups scored significantly higher than retention criterion groups. Within the retention groups, the one day feedback delay group achieved the highest mean score while the immediate feedback group achieved the lowest. Differences between the highest and lowest means were statistically significant for both articles.

The results support the hypothesis that delay of feedback in the classroom may improve retention. The results also suggest that, with grade eight students, the optimal feedback delay may be about one day.

The significant interactions suggest different factors in acquisition and retention. This is possibly a partial explanation for the apparent contradiction between the generally accepted principle of immediacy of feedback and the results of the present study.

Session 12.11

Division E

THE IMPACT OF PERSONAL AND ENVIRONMENT FACTORS ON SCHOOL RELATED PERFORMANCE

Janet R. Brown, Queens College, Chairman

AN EXAMINATION OF SELECTED ACHIEVEMENT, ATTITUDE AND BACKGROUND VARIABLES IN DISADVANTAGED ADOLESCENTS, *Janet R. Brown*, Queens College of the City University of New York, and *Elyse S. Fleming*, Case Western Reserve University

The purpose of the study was to explore the interrelationships among selected achievement, attitude, and background variables in a group of disadvantaged adolescents. This study was conducted as a prior step to the evaluation of a special summer enrichment program sponsored by the Phillips Exeter Academy for the youngsters who constituted the sample for this study. Thirty-three black and white eighth grade boys and girls were chosen from six cities to participate in the Special Urban Program (SPUR), specifically designed for moderately achieving students. The children were slightly above average in IQ and reading level but came from economically poor home backgrounds.

The data consisted of self-concept, intelligence, reading achievement and writing achievement scores, grade point averages, academic aspirations, achievement motivation, and several background variables—social class, number of siblings, perceived parental authority patterns, and participation in school activities. The data were summarized in three correlation matrices (boys, girls, and total group), from which clusters of variables having similar correlational profiles were derived according to Tryon's method of cluster analysis.

The results of the analyses yielded two major clusters of variables—Cluster I consisting of the self-concept measures and Cluster II consisting of IQ, reading, writing, and social class rank. Self-concept was found to correlate with social class rank, number of siblings, and perception of parental authority patterns, but not with the more usual achievement variables. The achievement and intellectual variables were inter-correlated for this sample, with the exception of grade point average which, in fact, was correlated with no other variables when the boys and girls were considered separately. Academic aspiration and achievement motivation appeared to reflect different role expectations for boys and for girls.

The conclusions of the study call into serious question current assumptions about 1) the nature of school achievement for inner-city youngsters, 2) the nature of educational programs for disadvantaged youth, and 3) the means of evaluating outcomes of educational programs for the disadvantaged.

ENVIRONMENTAL FACTORS, STUDENT VARIABLES, POST-HIGH SCHOOL EMPLOYMENT ADJUSTMENT, AND POST-HIGH SCHOOL EDUCATION OF MALE NEGROES FROM THE PROJECT TALENT SAMPLE, *David E. Kapel*, Glassboro State College (Glassboro, N.J.)

Data from twelfth graders in the 1960 Project TALENT testing and a follow-up study conducted in 1965 were used to compare the socioeconomic environment, the academic aptitude, employment-related factors, and post-high school education of male Negroes. The focus of the study was to determine the effects of geographic region, community type, and the percent of blacks in a school on the variables used. The following null hypotheses were tested: (1) environmental-parameter groups cannot be distinguished in terms of post-high school employment adjustment and student factors; (2) there are no significant differences among the groups related to environmental factors; and (3) there are no significant selected environmental factors that influenced students in the type of post-high school education acquired and future educational plans.

Six environmental parameter groups were identified and defined (N = 224). A partial canonical discriminant analysis was used to test null hypothesis 1. The assumption was that environmental effects could be identified if either the groups were unique to each other or sets of groups were unique. The analysis also identified the criterion variables that were most significant in discriminating among groups. A multivariate analysis of variance was used to test null hypothesis 2. Chi-square tests were used for hypothesis 3. It was assumed that if a pattern of significant chi-squares appeared when an isolated environmental factor was studied, the isolated environmental factor under study had a significant effect.

Analyses of the data indicated that there were regional differences and, to a lesser extent, there were community differences. Negro density did not appear to have an effect on the variables tested. Hypotheses 1 and 2 were rejected, while hypothesis 3 was not. This study indicated that further investigation of regional and community differences might be more profitable than focusing on the effects of the school's racial composition on black students.

SOCIAL CLASS DIFFERENCES IN ANXIETY OF ELEMENTARY SCHOOL CHILDREN, *Thomas H. Hawkes*, Temple University, and *Robert H. Koff*, Stanford University

The major purpose of this study was to examine the effects of social class, sex, and grade on the responses of children to a general anxiety questionnaire composed of items from the Children's Manifest Anxiety Scale CMAS (Castaneda, McCandless and Palermo, 1956) and the General Anxiety Scale for Children GASC (Sarason, Davidson, Lighthall, Waite, and Ruebush, 1960). Through a content analysis of the test items, a secondary purpose was to explore theoretical and practical issues relevant to the manifestation of anxiety in culturally disadvantaged children.

A general anxiety questionnaire composed of 31 items from the CMAS and 9 items from the GASC was administered to 460 Ss. In addition, the 10 items of the Lie scale of the GASC were also administered to the Ss. 211 of the Ss were fifth and sixth grade students at a private, laboratory school associated with a large midwestern urban university. This school draws its pupils from middle and upper-middle class and predominantly white families. 249 of the Ss were fifth and sixth grade students in an inner city public school which draws its pupils from predominantly lower class black families. This school has a thirty percent turnover rate in school population.

There were no significant differences on anxiety with respect to sex or grade. There was a significant difference ($p < .01$) with respect to social class on anxiety. Over eighty percent of the Ss in the inner city school were above the mean on anxiety found in the private school. There were no differences with respect to grade, sex, or social class on the Lie scale.

Through a content analysis of the items, speculations were made concerning the theoretical distinction between fear and anxiety, types of anxiety (social vs. security) and their relevance to social class, and the implications for the need of specific theories and strategies for teaching the culturally disadvantaged student.

NEED ACHIEVEMENT, TEST ANXIETY, AND SELF-CONCEPT OF ABILITY AS PREDICTORS OF COGNITIVE PERFORMANCE OF SEGREGATED BLACKS, DESEGREGATED BLACKS, AND WHITES, *Berj Harootunian*, Syracuse University

This study investigated the performance of segregated black, desegregated black and white ninth-grade students on tests eliciting a number of cognitive variables and on measures of need achievement, test anxiety, and self-concept of ability. The major question of concern was whether the relationships between the variables were the same for the different groups of students. A number of empirical studies have reported that personality variables have differential predictive power for the cognitive or academic performance of black and white students, while other studies have indicated no such differences.

In this study, the dependent variables were the cognitive constructs identified as ideational fluency, word fluency, problem recognition, and judgment. The predictor variables were, in turn, need achievement, test anxiety, and self-concept of ability. The tests eliciting these constructs were administered to black students who were in attendance in either a racially isolated school or a racially mixed school. In addition, data were collected on the white students in school attendance with the latter black sample. These three groups were further classified by sex, yielding six groups for the data analysis. The number of students in each were as follows: segregated black males, 41; segregated black females, 47; desegregated black males, 31; desegregated black females, 26; white males, 313; and white females, 311.

The coefficients of correlation between need achievement, as measured by the French Test of Insight, and the six cognitive measures were the most significant and most consistent across the different groups. The only group for whom need achievement was not a relevant predictor variable was the desegregated black female sample.

For the segregated black males and females and the desegregated black males, no significant coefficients of correlation were found between test anxiety, as measured by the Test Anxiety questionnaire, and the cognitive variables; for the desegregated black females, one coefficient was significant; for the white males, two coefficients were significant; and for the white females, four were significant.

Self-concept of school ability, as measured by Brookover's scale, was significantly related to the cognitive performance of segregated black males in no cases and of desegregated black females in only one instance. For all of the other samples, self-concept of ability was related significantly to at least four of the six cognitive variables.

Possible reasons for the occurrence of the different relationships across the various groups are discussed and are considered in terms of their theoretical implications.

Session 12.12 **Division D**

EXPERIMENTAL DESIGN AND ANALYSIS

Jerry Doppelt, The Psychological Corporation, Chairman

MAXIMIZATION OF GENERALIZABILITY UNDER COST CONSTRAINTS, *Russell A. Chadbourn*, University of Illinois

This study of the stability of psychometric measures examines the procedures by which the choice of conditions for a decision study can be made so that maximal stability can be achieved within the resource limits of the investigation. The generalizability coefficient, defined as the ratio of the universe score variance to the expected observed score variance, is shown as a function of the number of levels used for each facet of the study. Equations are derived which enable the investigator to determine the optimum choice of number of levels per facet to obtain maximum generalizability under the constraint of various cost functions. General cost functions which are considered for two-facet studies include power, logarithmic, and exponential relationships between cost and levels of each facet.

Maximization of generalizability under these cost functions is illustrated with examples of its utility in varying psychometric settings. One such example considers generalizability of supervisor's ratings of student teachers over an infinite population of supervisors (J) and occasions (K). With a total budgeted cost of $325 per student teacher, the cost function $C = bJ + mK + d$ is utilized to find the optimum number of supervisors and occasions for maximum generalizability within the cost limits of the investigation. For a particular set of values of the parameters, the optimum of five raters results in a G of .742, while use of six raters reduces G to .453.

Examination of the results of this and other examples indicates that considerable gain in generalizability can often be achieved at little or no cost by use of these techniques prior to conducting a decision study.

MEASURES OF ASSOCIATION IN COMPARATIVE EXPERIMENTS: THEIR DEVELOPMENT AND INTERPRETATION, *Gene V. Glass* and *A. Ralph Hakstian*, University of Colorado

To address practical rather than merely statistical significance, Hays (1963) introduced omega squared (ω^2), a measure of *strength of relationship* between a categorical independent variable X and a continuous dependent variable Y. Kelley's (1935) ϵ^2, a similar but earlier statistic, was visualized (incorrectly) as a correction for bias in η^2, the correlation ratio. Peters and Van Voorhis (1940) recommended using ϵ^2 in comparative experiments for precisely the same purpose for which Hays recommended ω^2 over twenty years later. In practice, values for ϵ^2 and ω^2 will not differ appreciably.

Difficulties with the use of ϵ^2 (or ω^2) stem from the conceptual complexity of categorical independent variables such as "method of instruction." Such molar variables cannot be represented in their entirety by the levels actually included in an experiment. By adding a single level (e.g., a control group) the value of ϵ^2 (or ω^2) can change radically, misleading the researcher. The arguments presented against the use of ω^2 parallel, but elaborate upon, objections which Fisher raised against η^2 and Lindquist against ϵ^2.

Even when the variable is molecular, use of ϵ^2 (or ω^2) is questionable in the fixed-effects situation; a measure of association between two variables is interpretable only when levels of both are randomly representative. (Of course, in the random-effects model, ρ_I, the intraclass correlation coefficient, is appropriate.)

In a fixed-effects model, measures of strength of association are not appropriate. Instead, depending on the nature of the fixed independent variable, the researcher should use graphical analyses or such inferential techniques as multiple comparisons and trend analysis.

A BETA INDEX TO CONFIRM CAUSAL DIRECTIONS IN A CLOSED SYSTEM OF FIVE VARIABLES, *George A. Nigro*, Boston College

The underlying intuitive principle of this study is that the beta coefficient of an intermediate variable in a causal direction remains relatively constant as other system variables are introduced and controlled in stepped regression, whereas that in the acausal direction changes noticeably. The main strategy was to compare changes in beta coefficients of the intermediate variable from pairs of stepped equations in the causal direction to changes in the acausal direction of 3-variable paths.

200x5 normalized random numbers were generated and substituted in interdependent equations to produce 5 scores for each of 200 pseudo-people. Stepped-regression analysis was then applied on all possible 3-variable paths. Beta differentials on a given intermediate variable were computed and compared for opposite directions; the smaller of the two was treated as that of the causal direction and became the numerator of a beta index, and the larger became the denominator. The index persisted to obtain values between 0 and +1 in the 1 direction as other system variables were stepped in; the closer to 0, the stronger the causal direction.

But some remaining 3-variable paths imitated the known causal paths. Hence, further study was conducted to find support for the known causal paths and to eliminate the imitations. Comparisons of beta coefficients for end variables in regression equations and comparisons of correlations among the 3 path variables of known causal paths yielded mathematically consistent like orders of magnitude as the beta differentials.

It was concluded that the beta index, found wanting initially, had greater utility as a confirming device by the addition of the two modifications, in that all but the originally designed causal paths were eliminated.

TIME AS A VARIABLE IN EXPERIMENTAL DESIGN, *Richard Wolf*, Columbia University

In the typical true experiment, individuals are randomly assigned to two or more groups, treatments are assigned at random to groups, and a period of instruction occurs. Students are tested at the conclusion of the treatment period with one or more measures, and the resulting data are usually analyzed by an analysis of variance procedure. Appropriate conclusions are then made from the results.

Two prominent characteristics of this design and its variants are that there is a fixed period of instruction and variable performance of subjects on the outcome measures. The design proposed in this paper is different from conventional ones in that variable time and fixed standards of performance are employed. In the new design, individuals are still randomly assigned to groups, and treatments are assigned at random to groups. However, each individual spends as much time as needed to reach a specified criterion level. Measures of amount of time spent in learning are obtained and analyzed.

There are two chief advantages to such a design. First the measurement of the subjects' performance is greatly simplified. Instead of trying to measure performance along an equal interval scale, one need only determine whether subjects have met the specified criterion level. A second advantage is that time is measured on a ratio scale. Thus, it is possible to make ratio statements about treatments. Difficulties in executing such a design inhere in the specification of the criterion level of performance and in maintaining a record of time spent in learning for each subject. An example of application of the design to the training of anesthesiologists is presented and discussed.

THE POSSIBILITY OF PRETEST - INSTRUCTION INTERACTION IN THE CLASSROOM UNDER TWO CONDITIONS OF PRETESTING, *William B. Ware*, University of Florida; *Norman D. Bowers*, Northwestern University

This study investigated the possibility of pretest-instruction interactions occurring in a classroom setting. The analysis of the data utilized a 2x3 factorial design and test information collected from 169 *S*s in 8 classrooms. The first factor was subdivided into two levels of instruction, Instruction and No Instruction. The three levels of the second factor were Pretest Identical to Posttest, Pretest Equivalent to Posttest, and Posttest Only. Tsao's general solution for the analysis of covariance was employed, and planned comparisons were completed to test the null hypotheses.

The results of the analyses indicated that a pretest-instruction interaction did not occur when the pretest was identical to the posttest, nor when the pretest was an equivalent form of the posttest. The results also indicated that *S*s who had been pretested did not perform significantly differently from *S*s who had not been pretested, both when the pretest was identical to the posttest and when the pretest was an equivalent form of the posttest.

Three possible explanations can be offered for these results. These explanations relate to the pretest situation, the passage of time, and the type of questions contained on the pretest. (1) If the pretest situation is similar to the situation to which one wishes to generalize, pretesting may not be associated with the occurrence of a pretest-instruction interaction. (2) If the time between the pretest and exposure to relevant instruction is long (four weeks in this instance), the pretested *S*s may fail to recognize the answers when they appear. (3) If the pretest questions demand the application of abstract principles to concrete situations, the occurrence of a pretest-instruction interaction may be less likely than if the pretest questions require the retention of factual materials. These explanations offered are hypotheses which must be examined before conclusions can be reached regarding the occurrence of pretest-instruction interactions in the classroom.

A COMPARISON OF THE EFFECTS OF TWO METHODS OF BLOCK FORMATION ON DESIGN PRECISION, *Louis A. Pingel*, University of Pittsburgh

For an experimental design consisting of B blocks, T treatments, and n experimental units per block-treatment combination, there are two possible methods of forming blocks when the blocking variable is measured on an interval scale:

(1) Predefined Range Blocking Method

Specify B mutually exclusive ranges in the population distribution of the blocking variable such that the probability of obtaining a value of the blocking variable in each range is 1/B, and then, for each of these B ranges, randomly sample Tn experimental units, all of which have values on the blocking variable in that range.

(2) Sampled Range Blocking Method

Randomly sample BTn experimental units, rank them according to their values on the blocking variable, and then take the experimental units with ranks 1 to Tn as the first block, the experimental units with ranks Tn + 1 to 2Tn as the second block, and so on.

In this paper, it is shown that, under the assumption that the joint population distribution of the blocking variable and the dependent variable is bivariate normal, heterogeneity of variance of the dependent variable is produced by both methods of block formation. Then, the within-block variances of the dependent variable and the design precision as a function of those within-block variances are compared for the two methods of block formation.

MAIN EFFECTS WHICH ARE NOT DIRECTLY TESTABLE IN ANOVA, *Andrew C. Porter*, Michigan State University; *Louis A. Pingel*, University of Pittsburgh

The purpose of the present paper is to state and offer proof of a theorem which facilitates the identification of main effects in a design for which the analysis of variance (ANOVA) does not provide an exact test of significance, even though ANOVA assumptions are met. Considerations are limited to designs where the independent variables are either fixed, finite, or random and where an estimate of within cell variation is available.

Let N denote the number of independent variables and r denote the number of these which are finite or random. If N is less than three, there is an exact test for every main effect. Consider any design where N is greater than or equal to three. Choose three of the N such that they are crossed with each other. They may be nested within some set of independent variables outside the triad. If the number of random plus finite independent variables in the chosen triad is two, then the fixed effect does not have an exact test, i.e., it has no error term. If the number of random plus finite independent variables in the chosen triad is three, then none of the three main effects in that triad has an error term.

Using the above theorem, it is shown that, for a completely crossed design where r equals 2, none of the N-2 fixed main effects has an error term; and where r is equal to or greater than three, none of the N main effects has an error term. If nesting occurs in the design, the above theorem may be used for all crossed triads. However, there are designs where the availability of an error term for some main effects may still be in question. These special cases are also dealt with in the paper.

Session 12.13 **Division E**

RESEARCH ON STUDENT PERSONNEL SERVICES AND STUDENTS

Thelma M. Urbick, Western Michigan University, Chairman

SEX DIFFERENCES IN COLLEGE FRIENDSHIP PATTERNS, *Albert H. Gardner,* University of Maryland

The purpose of this study was to determine sex differences in need similarity and complementarity in friendship, with emphasis on a thorough analysis of complementarity. *S*s, 259 males and 297 females, were students in the sophomore level introductory psychology course at a private university.

An experimental group of mutual same-sex friends was formed with 47 pairs of males and 52 pairs of females. A similar number of non-friend pairs was randomly chosen as a control group. The Edwards Personal Preference Schedule (EPPS) scores and the duration (months) and intimacy (five-point scale) of friendship were also obtained from *S*s.

The following results were obtained:

Intimacy. A person of the same sex was named as best friend by 91% and 88%, respectively, of the 259 males and 297 females in the population, with the mean intimacy ratings being significantly greater for females.

Similarity. Significant positive intraclass correlation coefficents between the same needs were found in the female experimental group for Autonomy (.41), Succorance (.23), and Heterosexuality (.49). None of the same-need relationships was significant for the control females, nor between the male groups.

Complementarity. There was also a lack of complementary matching in opposite EPPS needs for males, but nine of these combinations were significantly correlated for the experimental females: Achievement—Dominance (—.28), Deference—Exhibition (—.24), Deference—Aggression (—.26), Order—Aggression (—.30), Exhibition—Endurance (—.29), Autonomy—Affiliation (—.28), Affiliation—Change (—.28), and Nurturance—Change (—.23). Exhibition and Change were positively related (.27).

The greater intimacy of friendship and the tendencies toward both similarity and complementary matching among females suggested that they may seek to satisfy a greater variety of interpersonal needs through same-sex friendships than do males. It would appear that attention to sex differences should not be overlooked in investigations in this area.

COLLEGE STUDENT PARTICIPATION IN CONSTRUCTIVE SOCIAL ACTION, *Daniel C. Neale* and *David W. Johnson*, (both) University of Minnesota

The purposes of the investigation were to estimate the extent of constructive social action among college students and to identify characteristics that distinguish those heavily involved in such activity from college students generally.

A questionnaire was administered to 508 randomly selected students at the University of Minnesota and 100 volunteers in a tutorial project. On the basis of over 80% return, results indicated that less than 3% of the random sample were as involved in constructive social action as were volunteers. Volunteers and sample members were not significantly different on measures of social responsibility, knowledge of and participation in campus affairs, family social status, perceived value of social action activities, present source of income, or time commitments to study, work, and family. Volunteers, compared to sample members, were more likely to be older, to have higher grade point averages, to have majors in psychology or social sciences, to have come from small towns and suburbs, to live close to campus outside of family homes, and to report favorable opinions and high participation in social action on the part of models and reference groups.

The characteristics of student "constructivists," as identified in the study, were contrasted with findings from other research on student "activists." Results were also discussed in terms of student motives for constructive social action and implications for colleges and universities.

A STUDY OF STUDENT ACTIVISM, *Leonard L. Baird*, American College Testing Program

This study examines the characteristics of college students who show little, moderate, and great degrees of behavior of the sort called "student activism." Students are placed in these groupings according to their responses on a scale of student activism. Items on the student activism scale refer to attempts to change institutional rules, participation in demonstrations, working actively in an off-campus political campaign, etc. These data come from a follow-up of the American College Survey (ACS) (Abe, et al., 1965), administered to 2,295 men and 2,834 women college sophomores in 29 colleges. Data from the ACS were gathered when the students were freshmen and included interest test scores (Holland's VPI), measures of potentials for achievement, high school achievements, educational values, self-ratings, and life goals. The follow-up data included measures of college academic and nonacademic achievement and checklists of college experiences. One way analysis of variance will be used to examine the differences between the three groups of students on these variables. These data will allow us to examine the characteristics of student activists when they entered college, and to study their experiences and achievements during their college careers.

The results should provide a comprehensive portrait of the student activist and contrast him with students who are inactive or only occasionally active. The size of the sample and the diversity of institutions suggest that the results of this study will have some generality.

THE SCOPE OF ORGANIZED STUDENT PROTEST IN 1967-1968, *Richard E. Peterson,* Educational Testing Service

The first of two major purposes of this study was to describe the extent of organized college student protest in regard to specific issues during the past academic year, giving special attention to numbers of institutions (total and by type) affected and the proportions of student bodies involved. The second major purpose was to examine trends in student activism between 1965 and 1968 on the basis of comparable data gathered in 1965 (reported in the monograph entitled *The Scope of Organized Student Protest in 1964-1965*).

In late May of 1968, a machine readable questionnaire was sent to the dean of students or comparable official at all 1000 accredited four-year colleges in the country. The return rate was 86%. Respondents indicated for each of 27 issues (student participation in campus governance, the draft, military recruiters, etc.) if there had been protest and, if so, its extent and the proportion of students involved. Certain other information about the college (type, size, location, etc.) was also elicited.

Vietnam, dormitory rules, civil rights, and student participation in campus governance were the most widely protested issues, in that order. Matters related to instruction, curricula, faculty affairs, and freedom of expression were relatively infrequently at issue. Southern colleges less frequently experienced protest over the off-campus issues (the draft, etc.). Particularly for the off-campus issues, there were marked differences by institutional type: public and independent institutions reported more and sectarian and career oriented colleges (teachers, technical) reported less organized activism. In contrast to the 1965 data, student protest is generally more widespread: twice as many colleges reported Vietnam demonstrations, and interest in "student power" issues is on the rise. However, active concern about civil rights declined. Proportions of student bodies involved has *not* increased appreciably (although the absolute number of student activists probably has).

PERSONALITY CHARACTERISTICS, ENVIRONMENTAL PRESSES, AND ATTRITION FROM COLLEGE, *Robert Cope,* University of Massachusetts

Environmental characteristics at a large public university were found to be related to attrition. The assumption guiding the investigation was that a proportion of the dropouts from an institution of higher education can be attributed to a lack of "fit" between student characteristics and the social and academic presses of the institution.

Extensive social-psychological data were gathered by questionnaires and the Omnibus Personality Inventory on two entering freshmen classes ($N = 4150$). Two years later the students who had withdrawn ($N = 1131$)

from the University were contacted in a followup survey. The followup survey (80% response) determined why the students dropped out and sought to assess the nature of the problems they experienced while in attendance.

The social-psychological data on students who persisted were compared to the data on students who dropped out. A comparison of these data suggested that the salient environmental characteristics of the institution (large, liberal, affluent, secular, academically competitive, and cosmopolitan) were related to attrition, and that the relationship differed according to the sex of the student. Male students tended to drop out if they were more religious or politically conservative. Female students tended to drop out if they were less esthetically inclined, came from less wealthy homes, considered themselves less attractive, or had lower verbal aptitude test scores. Both males and females from smaller communities and those having lower mathematical aptitudes tended to drop out. Furthermore, the dropouts could be "typed" by the nature of their problems while in attendance. A factor analysis suggested that there were several types of dropouts: academic, social, religious, and too intellective.

The findings suggest that various college and university presses are related to student behavior, and that the same environmental press may have different effects depending on the sex and the social-psychological characteristics of the student. Also suggested by these data is the notion that the many reasons given for dropping out can be summarized by a typology of dropouts.

COMPARISON OF COLLEGE EXPECTATIONS AND EDUCATIONAL GOALS FOR FOUR COLLEGIATE STUDENT GROUPS, *Clarence H. Bagley*, State University of New York at Cortland

Educational Testing Service introduced in 1965 the Institutional Research Program in Higher Education, with two primary testing instruments: the College and University Environment Scales (CUES) and College Student Questionnaire (CSQ). The CUES is an adaptation of the earlier CCI instrument by Pace and Stern, while the CSQ was developed for assessing attitudinal as well as factual data on student characteristics.

Both the CUES and CSQ have been used in the Four College Study of Institutional Development within the State University of New York. The four colleges use the CUES and CSQ instruments to provide basic descriptive data concerning their students.

The study was designed to answer three general questions:

(1) What perceptions of the institution were held by each of six described groups, and to what extent and when does the "halo" effect of expectation by freshmen on the CUES diminish to a "reality" level as held by the upperclass students or faculty?

(2) When comparisons are made for the four different student groups (Clark-Trow's typology of vocational, academic, collegiate, and nonconformist) as found in the CSQ, what perceptions (CUES scales) do these four groups have and what changes are made from the freshman to sophomore year?

(3) Do item responses on the CSQ regarding level of proposed academic training and educational level of parents—both primary indicators of intellectual goals and choice of major—compare with CUES scale scores as well as the CSQ typology of student group?

The data for the study were taken from students in school during the 1966 and 1967 academic years at the four colleges. The sample groups were identified as (1) 1966 freshmen, (2) 1966 sophomores, (3) 1966 juniors, (4) 1966 seniors, (5) 1966 faculty, and (6) the 1966 freshmen tested in 1967 as sophomores.

Analysis of the data showed:

(1) The "halo" freshmen expectations on the CUES are near reality by the middle of the second semester, and that the sophomore-senior scale scores are very similar to faculty scores. Thus the "halo" expectations for most freshmen diminish rapidly during the latter half of the first semester of school and reach a plateau during the second semester.

(2) The four student groups were only moderately correlated with the CUES scale scores. Tho four groups showed little differentiation among themselves when compared on the CUES scale scores, reflecting the homogeneity of the student body at these institutions.

(3) Student academic aspiration level on the CSQ shows a more significant correlation to the CUES scale scores than did the student typology classifications on the CSQ. There was also a greater differentiation among classifications of student academic aspiration level when compared on CUES scale scores.

DELAYED AND NORMAL PROGRESS COLLEGE STUDENTS: A COMPARISON OF PSYCHO-SOCIAL CHARACTERISTICS AND CAREER PLANS, *Lyle F. Schoenfeldt*, University of Georgia, *Allen E. Bayer*, American Council on Education, *Marsha D. Brown*, American Institutes for Research

Studies of student progress through undergraduate institutions have traditionally focused on comparison of dropouts to those exhibiting the "normal" pattern of college completion. The purpose of the present study is to consider a group that has been generally ignored, those students who delay entry into college.

The basic sample included participants in the nationwide Project TALENT 5% probability sample who were 11th graders in 1960 and had responded to the 1962 and 1966 follow-up questionnaires. Field interviewing of a random sample of nonrespondents made it possible to correct for respondent bias.

Three groups were identified: (1) those who reported *normal progress* and receipt of a baccalaureate within the standard 4-5 years after high school, (2) those who *delayed entrance* to college but reported substantial progress toward a degree by 1966, and (3) a *control group* who acquired some post-high school education (college or noncollege) but were not enrolled in college and had not received a baccalaureate in 1966. These groups were compared on maximum performance scores, noncognitive scales, socioeconomic background items, and follow-up data which constituted the Project TALENT files. Differences between men and women were also analyzed.

Well under one-half of college entrants followed a "normal" pattern. Of those who did not, a substantive proportion delayed entry. Few of the "normals" or "delays" married before college as compared with those who acquired other types of education. In general the "delays" obtained jobs immediately after high school, although a substantial proportion of the males entered military service. In terms of academic aptitude, the "delays" of both sexes were similar to the control group but differed from the first group of "normal" college graduates. However, the "delays" did not differ from the normal progression group in terms of socioeconomic background.

Delay of college entrance is discussed as providing a valuable interim period during which an individual can "mark time" while reassessing personal educational values and crystallizing new motivation for educational achievement.

STUDENT TYPES AND THEIR ATTITUDES TOWARDS SELECTED SOCIAL, ECONOMIC AND POLITICAL ISSUES, *Irvin J. Lehmann* and *Charles G. Eberly*, Michigan State University

For some time now, but especially in the last few years, there has been growing concern among faculty members concerning the attitudes and values possessed by college freshmen. No doubt, the uprisings at Berkeley, Columbia, and Wisconsin, to name a few, have added to this concern.

The purpose of this research was to ascertain whether there was a marked difference in the attitudes and values shared by students who, a priori, were classified into one of four categories: academic, vocational, social, and non-conformist.

In the course of a larger study concerned with the characteristics of an entering freshman class, a detailed questionnaire was administered to about 6,500 freshmen entering a large, midwestern state university. Approximately 50 items of this questionnaire were concerned with ascertaining the attitudes of these students to such issues as "Red China should be admitted to the U.N.", "Abortion should be legalized for certain reasons", and "Marijuana should be legalized". The students responded to each item according to a five-element key (strongly agree to strongly disagree; the mid-point was a 'no-opinion' stance).

Because of the few numbers at the extremes, the cells were collapsed to yield only an "agree-neutral-disagree" continuum. These data were then analyzed to ascertain whether there was any relationship between the student's typology and his manner of responding to each of the items. Those items that exhibited significant differences were identified and explanations were offered to account for the variation in responses.

ROLE ORIENTATIONS TOWARD COLLEGE AS PREDICTORS OF OVER AND UNDERACHIEVEMENT, *Harry Schumer*, University of Wisconsin-Madison and *Victor Savicki* and *Robert E. Stanfield*, University of Massachusetts

The purpose of this study was to determine a profile of role orientations and scholastic ability of incoming freshmen that would best describe and discriminate between overachievers, achievers, and underachievers. Role orientations were measured by the eight factor scores of College Behavior Preference Schedule (Vocational, Academic, Intellectual, Consummatory Collegiate, Instrumental Collegiate, Social Development, Ritualistic, Greek). Scholastic ability was defined by the Predicted Grade Point Average obtained from the university admissions office. The Achievement categories were defined at the end of two semesters in the following manner: *Overachievers*—Grade Point Average (G.P.A.) more than .5 above Predicted

Grade Point Average (P.G.P.A.); *Achievers*—G.P.A. within ± .5 of P.G.P.A.; *Underachievers*—G.P.A. more than .5 below P.G.P.A.

A multiple discriminate analysis yielded two discriminate functions ($p < .004$). In the first function ($p < .006$) underachievers were distinguished from achievers and overachievers (multiple range, $p < .01$) and received positive loading on P.G.P.A., Greek and intellectual dimensions, and negative loadings on Ritualistic, Vocational and Social development dimensions. On the second function ($p < .11$) the overachievers were significantly different from the achievers, but neither were different from the underachievers ($p < .01$). The overachievers received positive loadings on the Academic, Vocational, Social development and P.G.P.A. dimensions, and negative loadings on the Greek and Consummatory collegiate dimensions.

The results indicate that underachievers tend to be high in scholastic ability, strongly oriented toward fraternity and sorority activities, interested in intellectual and artistic pursuits, and quite independent of their families. In addition, underachievers tend not to view college as a means to some future occupational goal, nor are they particularly concerned or interested in people. Overachievers, as distinguished from achievers, seem to have strong preferences for academic behaviors and vocational oriented activities, moderately high scholastic ability, an interest in meeting and serving people, and a dislike for hedonistic collegiate behaviors.

The results tend to point up the importance of "non-intellective" factors (i.e. role orientations) in predicting academic achievement and nonachievement.

CURRICULAR CHANGERS AND PERSISTERS: HOW DO THEY DIFFER? *Robert D. Brown*, University of Nebraska

The purpose of this investigation was to determine whether there are identifiable characteristics that distinguish college students who change their academic goals from those who persist. Little is known about whether changers and persisters differ as a group or whether they differ within particular fields. Are changers different even at the start of their college career? Are curricular changes isolated events or are they accompanied by changes in attitudes and personality?

Two hundred and forty-four male college freshmen, 122 science and 122 humanities students, were sampled from a larger group of changers and persisters. Persisters, 61 in each curriculum, were enrolled in the same academic major at the end of the year, while changers, 61 in each curriculum, had shifted their major to an unrelated academic field. The *Omnibus Personality Inventory* and an activities check list were administered at the beginning and end of the year.

Two-factor analysis of covariance, permitting testing the interaction effect of change and curriculum as well as the main effects, was employed for analysis of end of the year results.

There were significant differences between changers and persisters on attitude and personality measures at the beginning and end of year with changes occurring during the interval. Significant interaction effects revealed that changers became more like persisters in the other curricular group.

These results suggest that changes in academic goals are not isolated decisions of finding the proper vocational niche, but they can be part of a developmental process which includes changes in attitude and personality.

THE IDENTIFICATION AND EVALUATION OF COLLEGE EFFECTS ON STUDENT ACHIEVEMENT, *Donald A. Rock, John A. Centra* and *Robert L. Linn*, Educational Testing Service

The purpose of this study was to identify characteristics of colleges whose seniors have high academic achievement scores relative to their scholastic aptitude scores four years earlier. The identification of such characteristics could be of value to prospective college applicants and to administrators in making decisions about the allocations of resources.

The sample consisted of 84 colleges whose seniors had taken both the Graduate Record Examination (GRE) Area Tests in 1967-68 and the College Entrance Examination Board's Scholastic Aptitude Tests (SAT) in 1963-64. Seniors with GRE Area Test scores (Humanities, Natural Sciences, Social Sciences) in 1967-68 were randomly sampled within each college, with a maximum of 100 students selected from any one college. Colleges selected for the study had administered the GRE tests to essentially all seniors. SAT scores from the 1963-64 test administration were obtained for students in the sample, resulting in a total of 6160 students with usable test results. College means were then computed for the GRE and SAT scores.

College means based on the GRE Area Tests were used as output measures, and SAT means were used as input measures. College descriptive characteristics, such as income per student (I/S), number of books per student, faculty student ratio, proportion of faculty with doctorates (FD), number of students, and percentage of students in various major fields of study were then used to form taxonomic groups of colleges with different mean residuals obtained from regressing output measures on input measures.

Approximately 80 percent of the between-college GRE output variance was predictable from SAT input scores. College characteristics which accounted for a significant proportion of the 20 percent residual variance were I/S and FD, both being associated with positive mean residuals.

DIFFERENCES BETWEEN COLLEGE AND NON-COLLEGE UPWARD BOUND STUDENTS, *Byron R. Egeland*, Syracuse University and *David E. Hunt*, Ontario Institute of Education

Project Upward Bound (UB) is intended to generate the skills and motivation necessary for college success among high school students from low-income backgrounds and inadequate secondary school preparation. From June 1966 to June 1968, the Youth Development Center, Syracuse University, conducted a national evaluation of the effects of UB on students' attitudes and motivations (final report, Hunt and Hardt, 1967).

As part of the overall evaluation, the present study was conducted to determine whether or not UB students who went on to college differed from UB students who did not attend college. A sample of 1302 UB students from 21 different programs was given a battery of tests consisting of such measures as Importance of College Graduation, Possibility of College Graduation, Self-evaluated Intelligence, Motivation for College, Interpersonal Flexibility, Self Esteem, Internal Control, Future Orientation, and Non-alienation. These measures, plus other indices of change, were given during the first and last week of the eight-week intensive educational study program (UB summer program, 1966) and during the following spring (March 1967). Demographic information was obtained at the beginning of the summer program.

Counselors in the UB students' high schools were contacted, and information regarding whether or not the student went to college was obtained. Fifty-eight UB students who attended college were matched on high school GPA, UB program attended, and high school curriculum (e.g., college prep or vocational) with 58 UB students who did not attend college.

Using a 2 X 3 repeated measure ANOVA, the results show interactions and differences occurring on the measures of Interpersonal Flexibility, Future Orientation, Self-evaluated Intelligence, Importance of College Graduation, Possibility of College Graduation, and Motivation for College. The results show that students' attitudes changed during the summer UB program, and further changes also occurred during the following school year. These changes were related to whether or not the UB student attended college. In addition to the nine measures reported in this abstract, differences and trends occuring on other measures of change will be discussed. Demographic information will also be reported.

FACTORS RELATED TO PERSISTENCE AND WITHDRAWAL AMONG UNIVERSITY STUDENTS, *Jack E. Rossmann*, Macalester College and *Barbara A. Kirk*, University of California at Berkeley

In an attempt to understand attrition among university students, students who returned to the University of California at Berkeley for their sophomore year were compared with students who failed to return. During the week preceding registration for their freshman year, 1099 men and 1106 women who were planning to enroll in the College of Letters and Science were asked to complete the School and College Abilities Test (Form UA), the Omnibus Personality Inventory, and a biographical questionnaire. Over 80 per cent of the women and 60 per cent of the men completed all the instruments. (Over 90 per cent of both groups completed the SCAT.)

The students were categorized as persisters (N = 1713), voluntary withdrawals (N = 214), or failures (N = 258). Data analyses indicated that the withdrawals (students who had 2.00 g.p.a.'s or better at the end of their freshman year but failed to return for their sophomore year) had significantly higher SCAT-verbal scores than the persisters (students who had 2.00's or better at the end of their freshman year and returned for their sophomore year.

These two groups also differed ($p < .05$) on 7 of the 14 OPI scales. These differences indicated that the withdrawals were more likely to enjoy reflective or abstract thinking, were more interested in artistic activities, tended to be more tolerant of ambiguities and uncertainties, were more ready to express their impulses, were more intellectually oriented, and were less interested in a practical or applied approach to life. Chi-square analyses of the questionnaire data indicated that the withdrawals were less conventional and somewhat more activist in orientation than were the persisters.

Comparison of the failures (students who had less than 2.00 g.p.a.'s when they withdrew from the Berkeley campus) and withdrawals suggests that the failures were significantly less able academically and less intellectually oriented.

The results lend support to those who contend that among withdrawals from universities today are many very able and creative students. The data also suggest that voluntary withdrawals and failing students are quite different groups and should not be pooled as "dropouts."

AN INVESTIGATION OF THE EFFECTS OF THE UNIVERSITY CLIMATE ON THE IMPULSE LIFE OF THE STUDENT, *Thomas A. Lazzaro*, Veterans Administration Hospital, (Syracuse, N.Y.) and *Donald L. Beggs*, Southern Illinois University

This study was designed to investigate the hypothesis that movement from relatively restricted home and peer relations into a wider social environment during the college years results in a lessening of restrictions on impulse life.

The Self-Report Test of Impulse Control (STIC) is a single scale paper and pencil instrument developed to measure the reported impulse control of an individual. The instrument has been validated with respect to university populations and has been shown to be reliable over a short period of time ($r = .89$).

The STIC was administered to a freshmen sample drawn from six English classes. Due to the fact that age may have some effect on impulse control, no freshman over twenty years of age was included in this sample ($N = 96$). In addition, the STIC was administered to 137 university upperclassmen (ages 19-22) enrolled in four undergraduate educational psychology courses.

Although the testing procedures and examiners were identical in the two groups, many of the freshmen students were reluctant to take the scales while excellent rapport was maintained with the upperclassmen.

The results of this study indicate that the university climate does have some effect on the impulse control of an individual, for it was found that university upperclassmen have more control over their impulse life than university freshmen. This seems to contradict the findings of previous research studies that the university climate lessens the restrictions on the impulse life. The results of this study imply that the climate increases the restrictions on the impulse life of the student. Although several implications are obvious, the authors would suggest that freshmen today are quite different with respect to impulse control from the university freshmen five years ago. It is quite possible that the university climate today is being used by the student as a means to control his impulse life.

Session 13.5 **Division C**

COGNITIVE DEVELOPMENT: MULTIPLE CLASSIFICATION

Laurence D. Brown, Indiana University, Chairman

A DEVELOPMENTAL STUDY OF THE EFFECTS OF THE PERCEPTUAL SUPPORT ON CHILDREN'S ABILITY TO CLASSIFY, *Aurora L. Biamonte*, University of Illinois

This investigation checked the compatibility of Wohlwill's and Piaget's interpretations of mental development by exploring the role of stimulus variables in cognitive development. Wohlwill's propositions were the basis for a set of hypotheses that asserted that the attributes of stimuli differentially affected cognitive functioning at various age levels. The hypotheses based on Piaget's position indicated that mental structures provide the significant variable. Subjects were 96 lower-middle class girls selected from kindergarten, first, second, and third grades of two suburban schools. At three different times each girl was individually presented with a set of 16 Flagg dolls and asked to perform a series of cognitive tasks which tested classification, class inclusion, and conservation. Only the classification tasks will be discussed in this paper.

The attributes of the stimulus objects were dichotomously distributed in each set. Over the successive interviews the distribution of three of the attributes—hair color, shoe color, and dress pattern—were varied to correlate with age or to be independent of age. As a consequence, the redundancy and the noise in the stimulus objects present to each child differed over the three sittings. The order of presentation of the three sets of stimulus objects was counterbalanced within each grade level; the cognitive tasks were held constant across the three sittings.

The classification tasks were administered by Kohlberg's technique and scored in terms of:

1. the number of dolls included in a sort,
2. the type of dolls selected in a sort,
3. the verbal justification of the sort.

The initial analysis of the data indicates that, at each grade level and across the sets, the attributes of the dolls were not a major determinant of how many dolls were included in a sort. While the type of dolls selected was strongly influenced by the distribution of the attributes in a set, the reasons given for selecting the dolls at each grade level were stable across the sets.

The data from this experiment lend support to Piaget's structural theory of cognitive development and suggest directions for refining Wohlwill's propositions.

THE LEARNING, TRANSFER AND RETENTION OF DOUBLE-CLASSIFICATION SKILLS BY FIRST-GRADERS, *Paul I. Jacobs* and *Mary Vandeventer*, Educational Testing Service

The double-classification problem forms the basis of many intelligence test items, and is of major interest in Piagetian theory. The solution of such a problem requires the child to take into account simultaneously two different dimensions of stimulus variation (e.g., size and shape) while inferring a logical relation. The present study utilized highly structured yet individualized training techniques in an attempt to teach the skill of solving double-classification problems to first-grade children.

Following a pre-test of double classification skill with stimuli varying in color and shape, 42 first-grade *S*s were randomly assigned to either an experimental or control treatment. Experimental *S*s were individually instructed by *E* in the skill of solving double-classification problems with stimuli that varied in color and shape. Instruction was terminated when *S* reached a predetermined level of proficiency, or when a half hour elapsed. Control *S*s were given individual attention for an equivalent amount of time while engaged in an irrelevant task. Several hours following the treatment, each *S* was retested by a second *E* who had no knowledge of what group *S* was in. Both a post-test (a randomly parallel alternate form of the pre-test involving the dimensions of color and shape) and a transfer test involving the dimensions of shading and size, were administered.

Experimental *S*s scored significantly higher than control *S*s on both learning and transfer tests. *S*s were tested for retention four months later. Again experimental *S*s scored significantly higher than control *S*s on both learning and transfer tests. The implications for Piagetian theory and for intelligence test construction are discussed.

THE UTILIZATION OF CONCRETE, FUNCTIONAL, AND DESIGNATIVE CONCEPTS IN MULTIPLE CLASSIFICATION, *Ronald K. Parker* and *Mary Carol Holbrook,* Florida State University

A matrix task was presented to kindergarten, first, second, and third graders to investigate developmental changes in multiple classification. Correct solution of the incomplete matrices, comprised of three pictures in a row and three pictures in a column meeting at a blank intersection, required identification and combination of the common attributes of the row and the column. Three types of concepts (concrete, functional, and designative) were used in construction of the matrices.

Performance on the matrices improved with grade level, and there was a significant interaction between grade and type of matrix. Multiple classification of concrete concepts preceded the ability to combine functional concepts, with the ability to combine designative concepts appearing last. In general, when errors were made, children seemed to choose the picture representing the type of concept (concrete, functional, or designative) shown by object-sorting studies to be the most frequent mode of categorization. Latency measures decreased with practice and were negatively correlated with correct matrix solution. The consistency of the results with the theoretical positions of both Piaget and Bruner was discussed.

Session 13.6 **Division C**

INDIVIDUAL DIFFERENCES NO. 3

Merrill Sitko, University of Michigan, Chairman

PRE-SCHOOL MEASURES, SEX, AND GRADE ONE ACHIEVEMENT,

Robert Huebner, University of Maryland

Twelve pre-first grade measures were used in multiple regression and correlation analysis for the purpose of identifying variables that predict various criteria of success in grade one (Total N = 79, boys = 39, girls = 40).

The pre-measures were the six subtests of the *Metropolitan Readiness Test* (MRT), three parts of the *Maturity Level for School Entrance and Reading Readiness Tests* by Katherine Banham, a word learning list, the student's age, and whether or not he had had kindergarten experience. The criteria used were the four subtests of the *Metropolitan Achievement Test* (MAT) and reading level assessed at the end of grade one.

Several of the MRT subtests and chronological age made significant contributions toward predicting the various criteria. However, the Banham measures, the word learning test, and kindergarten experience made virtually no independent contribution toward predicting any of the criteria. MRT Numbers and MRT Matching consistently ranked one and two among the twelve predictors of reading performance for the total group. However, when the data were analyzed separately for boys and girls, the measures which were the best predictors for boys were among the poorest predictors for girls, and vice versa. Specifically, MRT Numbers was consistently the best predictor of each criterion for boys, but ranked at or near the bottom among the twelve variables in predicting the same criteria for girls. Differences in partial correlations were as striking as differences in rank. The partial *r* of MRT Numbers with the various criteria ranged from .43 to .50 for boys and from .02 to .04 for girls. Conversely, MRT Matching is the best single predictor of the various criteria for girls, but ranks low and contributes virtually nothing toward predicting those same criteria for boys. Equally striking sex differences were found when using other criteria and other predictors.

AUDING ACHIEVEMENT OF FIRST GRADERS AS RELATED TO SELECTED PUPIL CHARACTERISTICS, *Mary C. Nesbitt*, University of North Carolina at Charlotte

Auding is defined as the process of hearing, listening to, recognizing and interpreting or comprehending spoken language. Because of what is known concerning individual growth and development, one can expect to find in any classroom situation a wide range of auding achievement within any chronological age group. The purposes of this study were primarily to identify the auding levels of achievement of first grade children, and to explore the extent of relationship between the child's auding achievement and selected pupil characteristics.

Eighty first graders selected on basis of socioeconomic classification, sex, and chronological age were individually administered a test to determine auding achievement levels. Four graded selections of approximately one-hundred words each were used with a series of questions for each selection. Other tests used for comparison purposes included the *Metropolitan Readiness Tests* and the *California Short-Form Test of Mental Maturity.*

Findings indicated first grade children can understand orally presented materials above first grade reading level. Mean per cents of correct responses on subtests of auding instrument revealed that 56.0 per cent of the children could listen with understanding to material designed to be read at the second grade level, 49.7 per cent at third grade level, 36.0 per cent at fourth grade level, and 30.0 per cent at fifth grade level. Auding achievement test scores were found to correlate significantly ($p < .01$) with total test scores for *MRT* and *CTMM*. However, correlations between auding achievement and the non-language section of *CTMM* were found to be non-significant.

It was perhaps noteworthy that the differences between sex favored boys. This finding is counter to expectations in reading achievement at first grade level.

THE INFLUENCE OF CULTURAL FACTORS ON PERFORMANCE OF FIRST GRADE CHILDREN, *Richard Thiel*, Southwestern Cooperative Educational Laboratory, Albuquerque

Children matched with parents from three distinct sub-cultures, Navajo, Pueblo, and Spanish-American, were studied to determine the influence of cultural variables upon children's performance on standardized tests as a first step in understanding cultural rather than individual styles of learning.

The children from three ethnic groups were selected from public school and P.I.A. systems. Each child was given a two hour battery of psycho-linguistics (I.T.P.A.), psycho-metric and achievement tests by examiners with specialized training. Mothers of these children were interviewed on a set of face-sheet variables from questionnaires and with a wide variety of attitudinal measures. Multiple regression analyses and factor-analyses were performed to ascertain the relationship between degrees of acculturation, assimilation of the parents, and their children's performance scores. Native interviewers were trained by shaping techniques in order to present faithful translations of the questions in English. Coefficients of concordants were used to check the reliability of the original taped translations. The home environment and cultural background indices were predictor indices, while children's test scores were criterion variables.

Significant differences were found among all three groups of children's mean performance scores ($p \leqq .01$), and differences in terms of acculturational and assimilational degrees were determined for sets of background measures ($p \leqq .05$) ($p \leqq .01$).

Culturally different learning styles suggest radical curricular changes to accommodate divergent children. Improving the educational opportunities of divergent children requires parental orientation programs and combined parent-child pre-school training.

THE RELATIONSHIP OF AN INDIVIDUAL'S ABILITY TO CHANGE COGNITIVE STYLE AND ACADEMIC ACHIEVEMENT, *Pearline Peters Yeatts*, University of Georgia

During the last decade a number of researchers have demonstrated the significance of cognitive style in concept learning. Kagan and his associates have suggested that cognitive style is a trait which pervades cognitive and perceptual activity.

The purpose of this study was to examine the relationship of an individual's ability to change style and his academic performance on verbal and quantitative task. Cognitive style was operationally defined in terms of the individual's performance on the Kagan Cognitive Style Test (KCST). It was hypothesized that individuals not able to change cognitive style would score as high on verbal and quantitative scores as those individuals who could change.

The KCST was administered to 62 fourth and 59 sixth grade students representing various ability levels. Subjects were tested individually and were asked to put each task together in as many ways as possible. The individual was allowed 45 seconds for each task. The subject's first answer was defined as his cognitive style Preference. Each answer thereafter was analyzed as to the sameness of style. The California Academic

Achievement tests were used to determine the verbal and quantitative abilities of the subjects.

Results of the analysis indicated a positive correlation between one's ability to change cognitive styles and verbal and quantitative scores on academic achievement test. The results further suggest that quantitative scores are affected less by the inability of analytic thinkers to change style. Individuals who had no difficulty switching from one style to another were found to score high on verbal and quantitative tests. Individuals falling in the middle (changing about half the time) performed at about grade level, whereas those not able to change at all were significantly below grade level.

The results of this study seem to suggest that the rigidity of one's cognitive style may have a significant influence on the manner in which he learns.

FURTHER RESULTS ON CERTAIN BLOOD CHEMICAL VALUES AS PREDICTORS OF COLLEGE STUDENT PERFORMANCE, *Richard H. Lindeman*, Teachers College, Columbia University, *Richard E. Gordon* and *Katherine K. Gordon*, University of Florida

The purpose of this paper is two-fold. First, we report results of further studies of the relationship between certain physiological and biographical characteristics and academic performance of students at Wagner College in New York City, a population which has been studied for the past five years. Second, we compare findings on these U.S. students with those on a sample of Swedish students at the University of Stockholm.

The results of earlier studies were reported in a paper read at AERA in February, 1966. The additional results presented here are based on data on serum chemical levels, SAT scores, grade-point averages, and certain biographical and personal characteristics of 75 male and 135 female Wagner College students. We found that relationships between serum uric acid (SUA) and grade-point average (GPA) were modified by other characteristics of the individual, namely, his personal drive and motivation, reaction to environmental stresses, and academic aptitude. We conclude that SUA level is positively associated with levels of motivation and inner drive required to compete in college. Thus SUA level may be a useful predictor of academic achievement among students who possess the requisite academic aptitude.

Interest in studying Swedish students stemmed from our hypothesis that under reduced environmental stresses and competition the observed relationship between SUA and achievement would be less pronounced or non-existent. Since studies in Swedish universities are typically individually paced, we believed that students there would provide a suitable sample. Data were obtained from a sample of 47 female and 31 male first and second year psychology students at the University of Stockholm. Results supported our hypothesis; achievement and SUA were found not to be significantly related. We also found differences between Swedish and American students in SUA and cholesterol levels as well as in certain attitudinal and personal characteristics. The implications of these findings are also discussed in the paper.

Session 14.2 **Division B**

USE OF OBJECTIVES IN CURRICULUM DEVELOPMENT

Elliot Eisner, Stanford University, Chairman

A STUDY OF TYPES OF GOAL STATEMENTS AND THEIR USES IN A CURRICULUM DEVELOPMENT PROJECT, *Decker F. Walker*, Stanford University

The nationally financed projects that have been major agencies for formulating curriculum during the last decade offer an arena for developing and testing curriculum theory. Marsh, Wooton, and Hurd have described the historical and administrative aspects of three such projects, but no studies have appeared which describe the processes by which these groups formulate objectives, select learning experiences, and organize them. This study began with the question of how members of a national curriculum project directed their individual efforts toward common ends through the use of *goal statements*—statements which enable one to visualize aspects of the completed curriculum, its products, or its effects. While educational objectives have traditionally been recognized as serving this purpose, preliminary observations of this project suggested that other less well-known forms of goal statements are also used. The purposes of this study were to identify the major types of goal statements used in this project and to determine which of these types project members perceived to be most useful to them in their work.

Five types of goal statements were identified: AIMS, EXPLANATIONS, CONCEPTIONS, EXEMPLARY PRODUCTS, and PROCEDURES. A content analysis of the publications of the project director was conducted using these five categories. The goal statements located in this way were then used to construct an instrument which was administered to all project personnel to discover whether some types of goal statements were perceived to be more useful than others in translating curricular intentions into acceptable curriculum products. The instrument consists of 108 statements of nine kinds: AIMS, EXPLANATIONS, CONCEPTIONS, EXEMPLARY PRODUCTS, PROCEDURES, and combinations of AIMS with each of the other types. Each project participant rated each statement on a five point scale according to the extent to which he believed the statement would be useful in enabling a new member to function effectively in the project.

Results of the study indicate that a distinction can be dependably made among types of goal statements and that such a distinction might prove useful to both curriculum developers and curriculum theorists.

INSTRUMENTATION OF BLOOM'S AND KRATHWOHL'S TAXONOMIES FOR WRITING EDUCATIONAL OBJECTIVES, *Newton S. Metfessel* and *William B. Michael*, University of Southern California, and *Donald Kirsner*, System Development Corporation

The purpose of the paper was to show how specific behavioral objectives can be formulated within the hierarchy of the major levels and sublevels of the taxonomy of educational objectives in the cognitive domain as set forth by Bloom and his co-workers (1956) and in the affective domain as reported by Krathwohl and his collaborators (1964). Such a framework should furnish a helpful base around which behavioral statements of objectives can be formulated and applied in the school setting. An educational objective in the cognitive domain was defined as a description of the behavior of an individual (learner or examinee) in relation to his processing information embodied in subject matter, and that in the affective domain as roughly an internalization at various levels of emotional involvement of stimuli in a social context.

The investigators formulated the objective as consisting of two major components: (1) the behavioral component, involving primarily a cognitive or affective process expressed in the form of an infinitive, and (2) the subject-matter content or product in the cognitive domain as a direct object or the emotionally attractive (or unattractive) stimuli in the social environment in the affective domain as a direct object (usually nouns or gerunds describing environmental stimuli or activities in the environment being perceived).

The instrumentation process involves the citation at each level of the hierarchy within the cognitive or affective domain of both *key infinitives* and *key direct objects* that may be matched (permuted) to formulate objectives. The entries in the tables which run several pages constitute the results of the instrumentation effort.

Applications of the table to the formulation of objectives in educational practice will be illustrated and discussed.

BEHAVIORAL OBJECTIVES: HANDFULS OF CORN? *Elizabeth H. Brady*, California State College

This paper is concerned with four questions:

a) What effects on curriculum process have resulted from the current preoccupation with behavioral objectives?
b) Should specific terminal behaviors be precisely predictable in the education of people?
c) How feasible is the contention that greater efficiency and more extensive mastery of specified behaviors will result from instruction programmed by behavioral objectives?
d) Can curriculum theory enhance pre-service and in-service education of teachers?

The tendency pointed out by Ralph Linton for people to become preoccupied with one thing at a time during periods of rapid culture growth, as do chickens to whom a handful of corn is flung, seems borne out by the current emphasis on specifying objectives in behavioral terms while ignoring other aspects of the total curriculum process. Some consequences are suggested.

The usefulness of stating objectives behaviorally as opposed to using specific materials or teacher goals to organize instruction is acknowledged. Level of generalization and a rationale for keeping learning situations open-ended, allowing for successive restatements of objectives, are examined.

The relative inefficiency of some programmed approaches is argued, and the task of identifying "pay-off" content and learning experiences conceived in terms of multiple objectives is discussed.

The value of curriculum theory in preparing teachers to think about curricular and instructional decisions is argued. The responsibility of schools of education to prepare process specialists, as well as measurement specialists, is considered.

Session 14.5 **Division C**

COGNITIVE DEVELOPMENT: PIAGETIAN CONSERVATION NO. 1

Joseph M. Scandura, University of Pennsylvania, Chairman

SUCCESSFUL NUMBER CONSERVATION TRAINING, *Abraham H. Blum*, University of Wisconsin-Milwaukee, *Carolyn Adcock*, Boston University

This study investigates the efficacy of a training procedure designed to facilitate the attainment of number conservation.

Subjects were 43 children ranging in age from 5-7 to 8-3; one-half were entering the first grade, while the others had just completed first grade.

Presence or absence of conservation was determined by classical Piaget procedures. From the responses, subjects were designated as non-conservers (17), transitional conservers (5), or conservers (21). Non-conservers and transitional conservers were randomly assigned to a training or control group.

The training procedure differed from the pre- and post-conservation tasks in that cardboard spacers were used to determine distance between the objects. Controls performed a number matching task.

After training, all but one of the non-conservers attained number conservation. On the other hand, only one non-conserver in the control group "became" a conserver.

In a retest of the second graders three weeks after training, no changes were found in the performance of the control group. In the training group, the only changes were one conserver "lost" conservation ability and one conserver was found to have transitional conservation ability. Interesting secondary findings were also obtained. All transitional conservation subjects indicated conservation in contraction manipulation, but made non-conservation responses during extension manipulations.

Since this training was successful in inducing relatively enduring transition or full conservation, it would indicate that we have a theory and method for the acquisition of this concrete operation in both laboratory and classroom training.

A CONCEPT STABILITY TEST FOR SPECIFIC GRAVITY BASED UPON THE PRINCIPLE OF THE SMEDSLUND CONSERVATION OF WEIGHT TEST, *Henry P. Cole*, Eastern Regional Institute for Education

Smedslund designed a trick demonstration which appeared to show that weight changes occur following shape transformations. He reasoned that a subject who had attained the concept of conservation of weight in a stable form would not be "perceptually seduced" by the trick demonstration but would maintain the logical conservation response. This procedure was effective in differentiating the stability of the concept in a group of subjects, all of whom would have been rated as conservers by a conventional test.

The principle of the Smedslund test was used by the author to design an instrument to assess the stability of the concept of "specific gravity" as defined by Inhelder and Piaget. Their work indicates that four common dominant perceptual cues lead to formulations of irrelevant principles which frequently displace the relevant logical principle. The instrument designed for the study reported here—the Cole instrument—consists of a series of seven trick demonstrations which deliberately emphasize the cues leading to the common false principles. The ten-item Cole test was group administered by means of slides, tapes, and printed materials to seven groups of subjects (total N = 563). The groups varied in amount of training and ranged in age from seventh grader to college senior. The instrument was determined to be valid and reliable. KR20 coefficients ranged from .73 to .89.

It was reasoned that subjects who had conceptualized the principle in a stable form would resist the strong "perceptual seduction" effect of each demonstration and maintain the use of the logical principle. Significant differences ($p < .002$) in performance between groups of the same age but varying in training level confirmed this hypothesis.

Unlike the conventional instrument, the Cole test provides a rigorous assessment of the stability of the concept of "specific gravity." Instruments similar in design and intent might help assess more realistically the effectiveness of instructional methods.

FLEXIBILITY AND CONSERVATION ACCELERATION, *Andre D. J. Cote* and *Charles C. Anderson*, University of Alberta

The primary purpose of the present study was to examine the possibility of using information regarding relative standings on the dimension of flexibility to explain differential rates of conceptual advance in learning situations involving cognitive conflict. Specifically, the current investigation tested the hypothesis that success on Smedslund's cognitive-conflict training, designed to accelerate the acquisition of the conservation of substance, is directly related to a high standing on certain flexibility measures.

The measures used to derive a flexibility factor score for each *S* were 1) the Inhelder-Piaget Measure of Retroactive Flexibility, 2) the Inhelder-Piaget Measure of Anticipatory Flexibility, 3) Blum's (1959) Child Transition Test, 4) Zigler's (1966) Restructuring by Classification Task, 5) Uznadze's (1966) Haptic Modality Set Tasks, 6) the Children's Embedded Figures Test.

All *Ss* (N = 129, Mean Age = 75 months) were tested individually by the E. *Ss* adjudged to be highly flexible and who received cognitive conflict training acquired conservation significantly more often than the following groups: 1) *Ss* receiving similar training and adjudged to be less flexible ($p = .005$); 2) no training control *Ss* who were adjudged to be highly flexible ($p = .02$).

The results support the main hypothesis and offer some hope to those training agents attempting to account for acceleration within a strictly Piagetian conceptual framework. Further, a battery measuring flexibility has been tentatively established, an essential condition for the selection and evaluation of acceleration techniques suitable for individuals operating at various levels of flexibility.

A DEVELOPMENTAL STUDY OF NUMBER CONSERVATION ATTAINMENT AMONG DISADVANTAGED CHILDREN, *Ofelia Halasa*, Division of Research and Development, Cleveland Board of Education

Number conservation mastery occupies a central role in Piaget's theory of intellectual development. This study attempted to evaluate the validity of Piaget's theory by investigating the developmental attainment of number conservation among disadvantaged children. Certain aspects of Piaget's theory were explored: (1) understanding of Correspondence and Conservation operations, (2) relationship of two criteria of number conservation, Equivalent judgment and Symbolic (Conservation) reasoning, and (3) the understanding of transitive relational terms (more, less, and same) with Correspondence and Conservation operations.

The experimental design included two operational tasks (Correspondence and Conservation), four age groupings, two-level classifications of learning ability, and task difficulty.

*S*s were seen individually in two sessions, with a two week interval observed. Tests for understanding of more, less, and same and of Correspondence were administered during the first session. In the second session, two tasks of varying difficulty level were presented in a counter-balanced order to *S*, with each task consisting of Correspondence and Conservation tests. *S's* judgment and reasoning were asked following each operational test.

It was found that age and learning level were directly related to frequency of judgment of Equivalence and Symbolic reasoning responses. Developmental pace of number conservation differed qualitatively between High and Low Learning groups for either criterion, with differences increasing as one went up the age scale. Conservation responses to Correspondence and Conservation tests were significantly correlated and not affected by difficulty level. Understanding of more, less, and same was not sufficient for conservation.

Piaget's general formulations were confirmed. Those responsible for curriculum development in compensatory educational programs at pre-school, kindergarten and 1st grade should familiarize themselves with stages that inner-city children at different ages are maturationally capable of, as well as individual differences arising from one's learning ability. Understanding of stage-dependent conceptualization could provide a valuable scale for diagnosing where the child is, and it gives an idea of the child's present capability in terms of his understanding of arithmetic processes.

Session 14.6 **Division A**

THE LOCAL SOCIOECONOMIC CONTEXT OF EDUCATIONAL ADMINISTRATION

William Monohan, State University of Iowa, Chairman

SOCIAL AND ECONOMIC INEQUALITY AMONG SUBURBAN SCHOOL DISTRICTS: A LONGITUDINAL STUDY, *G. Alan Hickrod*, Illinois State University, *Cesar M. Sabulao*, Illinois State University

Social and economic data on the school districts of five large standard metropolitan statistical areas were assembled for two points in time, 1950 and 1960. This particular phase of the study focused upon differences among suburban school districts, rather than upon central city versus suburban contrasts. The problem was to determine trends through time in the dispersion of resources among these metropolitan school districts. Two types of data were collected. The first type was designated "human resources," and the second type was designated "fiscal characteristics." The human resource category included such items as percentage college educated, median family income, and an index expressing the occupational composition of the districts. The fiscal characteristics included such items as property valuations, expenditures per pupil, and tax rates. Fiscal data could be collected directly from cooperating departments of education in the five states. Human resource data was more difficult to construct, because federal census data had to be translated into school district terms by the very time consuming process of comparing census tract maps with school district maps. The variance and the coefficient of variation were used as statistical representations of equality and inequality.

The principal finding is that there is a tendency for school districts in the metropolitan areas studied to become more equal on fiscal characteristics with the passage of time, but less equal with regard to the human resource variables. This is seen most clearly with regard to the occupational index hand and the property valuation variable. This study lends empirical support to hypotheses proposed by Robert Havighurst regarding a tendency for school districts to become less alike in terms of their occupational composition. There is, however, considerable variation from one metropolitan area to another, and also a good deal of variation between variables within the two large classifications of "human resources" and "fiscal characteristics."

This phase of the study has implications for state aid formulae. In most states "equalization" aid is distributed on the basis of property valuations. However, this study suggests that for at least some metropolitan areas the school districts are becoming more alike with regard to property valuations. It is with regard to human resources that schools are becoming quite different from one another. It would appear that if state governments really do wish to pursue equity goals they will have to find some means of taking into consideration human resource differences in their grant-in-aid schemes.

The report concludes with a brief outline of activities to be undertaken in the next phases of the project, now designated "Metro" at Illinois State University.

THE ECONOMIC BASIS OF MULTILATERAL COLLECTIVE BARGAINING IN PUBLIC EDUCATION, *Michael H. Moskow*, Temple University; *Arnold H. Raphaelson*, Temple University; *Kenneth McLennan*, Temple University

This paper presents a theoretical model utilizing economic analysis to explain why collective bargaining in education tends to be multilateral in nature rather than the more traditional bilateral bargaining between employers and employee organizations in the private sector.

The model utilizes demand and supply theory from Classical economics to explain budget size, salary levels, and employment in a school district. In order to make the theory applicable to public education, emphasis is placed upon the combination of conflicting interests expressed through the political process as a source of community demand for education. This combining of conflicting interests contrasts sharply with the singularity of purposes that form the basis for demand in the private sector. In public education, all taxpayers are required to pay for the service, regardless of whether their children attend the public schools and thus benefit directly from the service. Prices do not perform their typical rationing function. The taxpayer cannot switch to another supplier without incurring tremendous costs. As a result, political pressure (either individually or through interest groups) is the major means available to taxpayers to influence the price and quality of the product. At times, school board members represent separately identifiable groups of citizens within the community.

The lack of reasonably priced alternative products and the emphasis on political pressure provide a unique environment for collective bargaining with teachers. Multilateral bargaining, either directly or indirectly, occurs frequently on salaries because they form a major part of the total school budget. Non-economic issues of social or political significance, such as decentralization and racial integration of teachers or students where methods or goals are subject to political conflict, are also likely to result in multilateral bargaining when they become topics of negotiations.

AN EXPLANATION OF VOTER BEHAVIOR IN SCHOOL DISTRICT TAX ELECTIONS, *Byron H. Marlowe,* Ohio Education Association

Most of the reported studies of voter behavior in school district tax and bond issue elections identify a causal relationship between certain demographic variables and the resultant voter behavior. This study, a random sample of registered voters covering seven local tax levy issues over three elections, reveals that the "fairness" or "equity" of the individual's property tax burden is the major variable which accounts for one's voting behavior. In this study, the gross income and property tax payment were separated into five ascending categories. The "yes" voters on property tax issues were renters or those with income categories larger than their tax category. "No" voters were those with an income category lower than their tax category. This relationship held for all classifications of the voter—occupation, age, income, education, parentage, party affiliation, religion, amount of tax, and area of residence.

The phenomenon of increased support of local tax issues with the increased voter turnout is also explained by a tax to income relationship. In Ohio, passage of tax issues correlates very strongly with the size of voter turnout. Annually a greater percentage of school issues are passed in the general than in the primary or special elections, and there is a four year cycle with more issues passed in the Presidential election year than Congressional or off-year elections. In this study, voter turnout of those with income categories below their tax category was consistently high. As a result, the "no" voter is overrepresented in the low turnout elections (specials and off-year, and to a lesser extent in the primary and Congressional year elections).

ATTITUDES OF THE RURAL POOR TOWARD EDUCATION, *Ernest A. Vargas,* West Virginia University; *Alvin R. Carter,* West Virginia University

Financial backing for educational progress is largely dependent upon local electoral support. The rural poor of Appalachia have often exercised their voting power to defeat measures designed to improve their educational systems. Consequently, in the Appalachia region, educational progress lags far behind the rest of the nation. This paper identifies the educational values of the rural poor and describes attitudes critical to support of the educational institution.

In 1967, an attitude survey was administered to a sample of 940 respondents in a rural county in southern West Virginia. One of the educational attitude scales elicited general educational values, and the other related specifically to the operation of the school system in that county. The population was categorized into two groups, poor (defined by economic deprivation: $3,000 for a family of four or appropriate ratio between family size and level of income) and nonpoor.

The results show the rural poor do not differ from the rural nonpoor on educational values; however, the rural poor consistently evaluate the educational system's operation and resource conditions more favorably than the nonpoor. Statistically significant differences were found in most of the indices measuring attitudes about the education system, but none were found with respect to educational values.

School administrators in areas where the rural poor constitute a large segment of the electorate should clearly distinguish between the value of education and the merit and difficulties of operating educational programs in their school system area. In addition, many difficulties school administrators encounter in urban school systems are due to attitudes brought by rural migrants to the city.

PROBLEM SITUATIONS ENCOUNTERED BY SCHOOL PRINCIPALS IN DIFFERENT SOCIOECONOMIC SETTINGS, *Ray Cross,* University of Minnesota; *Vernon Bennett,* University of Minnesota

A body of literature in administration has been concerned with the influence of situational factors on status leaders. Pursuant to that literature, this research investigated the differences between problems encountered by principals of schools in low socioeconomic communities and those encountered by principals in high socioeconomic communities.

Within a large metropolitan school system, five elementary schools located in low socioeconomic areas and five elementary schools located in high socioeconomic areas were identified by means of a community socioeconomic index. The problems encountered by the principal of each school during a two-day period were observed and classified according to a three-dimensional administrative problem taxonomy developed by David W. Darling. One dimension of the taxonomy required the classification of each problem according to *type of skill*—technical, human, or conceptual (Katz, 1955)—needed by the principal in handling the problem. The second dimension employed Chester I. Barnard's *origin of decision* categories—appellate, intermediary, and creative. The third dimension of the taxonomy was composed of four *administrative functions*—educational program, developing personnel, community relations, and managing the school (Livingson and Davies, 1955). Since each problem was classified simultaneously on all three dimensions, the 3 x 3 x 4 paradigm provided a thirty-six cell taxonomy.

Chi square tests revealed that no principal's problem distribution on any of the three dimensions differed significantly from the combined distribution for principals in the same type of socioeconomic setting. Chi square tests comparing the combined problems for principals in "low" settings with the combined problems for principals in "high" settings yielded significant differences ($p < .05$) for all three dimensions of the problem taxonomy. The modal problem type for principals in "low" settings required human skill, was appellate in origin, and was concerned with the management function. The modal problem type for principals in "high" settings required technical skill, was creative in origin, and was also concerned with the management function.

The results suggest that principals of schools in low socioeconomic communities operate in a problem milieu which is different from that of principals of schools in high socioeconomic communities.

Session 14.7 **Division C**

CONTRIBUTED PAPERS

Stephen L. Yelon, Michigan State University, Chairman

DISCRIMINATION REDUNDANCY OF WORDS AND DIGITS AND THE RATE OF PROCESSING INFORMATION, *James Bosco,* Western Michigan University

Numerous studies have been reported dealing with the problem of estimation of the information processing rate or the duration of the interval from the inception of the stimulus to its reception in the brain. These investigations have generally made use of the binary digit or bit in order to quantify the stimuli. Other investigations indicate that characteristics of the stimuli, in addition to the amount of information carried by the stimuli, have effect on the rate of processing information. Previous investigations have shown that processing is facilitated if the subject is told in advance the set of stimuli to be used. This study reports on investigations dealing with a characteristic of the set of stimuli which may effect the processing rate, i.e., the discrimination redundancy of the set. The basic question is: If the amount of information carried by stimuli is held constant, does increasing the discrimination redundancy of the set of stimuli decrease the amount of time required to process the stimuli?

Fifty undergraduate students were placed at random into five groups. The stimuli for the experiments were words and digits. There were three levels of discrimination redundancy for digits and two levels of discrimination redundancy for words. The stimuli used in each group carried two bits of information. The information processing rate was calculated using a procedure developed earlier by Travers and Bosco (1967). This procedure consists of the tachistoscopic presentation of two stimuli to the subject. The briefest interval between the two stimuli at which the first can be recognized is the threshold information processing rate.

Analyses of the data revealed that the discrimination redundancy of the stimuli has an effect on the information processing rate. The plottings of the data fall on straight lines, with increases in discrimination redundancy of sets resulting in decreases in the processing rate.

A MULTIVARIATE ANALYSIS OF CHANGE OF CLASSROOM GROUP STRUCTURE, *Leellen Brigman* and *Evan R. Powell,* The University of Georgia

Rentz and Olson (1968) have shown the practicability of using Osgood's Semantic Differential to rate persons as concepts, yielding a description of group structure. The purpose of this investigation was to empirically describe changes in the structure of small peer groups and to relate these changes to the phase "theory" of Bennis and Shepard (1956).

Twenty-nine (29) subjects in two sensitivity training groups rated each other weekly for four weeks on a form of the Semantic Differential (SD) using twelve bipolar adjectives separated by a seven point scale.

Inter-correlations were obtained among the twelve SD scales and the resulting inter-correlation matrix (with unities throughout the principal diagonal) was factor analyzed by the principal axis method, retaining for rotation those factors with eigenvalues greater than unity. Rotations were accomplished by the Varimax procedure. Subsequently, factor scores were calculated for each individual as a person-concept rated by others, by week and by group. Osgood's three factor solution—Evaluation, Potency and Activity (EPA) was obtained.

Plots were generated, showing location in "semantic space" of each group each week for each pair of factors (EP, PA, EA). Changes were analyzed, related to "theory" and controlled for Dogmatism (Rokeach, 1958).

Results showed changes in both groups on the three dimensions: Evaluation and Activity decreased; Potency increased. Interpretations of these changes in terms of "theory" were offered. Implications of the use of this technique to describe changes in the affective dimensions of small peer groups were stated. These implications included the necessity for definition of group structure as a mediator of the presage-product relationship in the classroom.

A COMPARATIVE STUDY OF PRESCHOOL EDUCATIONAL TV PROGRAMS, *Loyd H. Inglis, Edward L. Palmer, Jack Crawford* and *Cathy Kielsmeier,* Teaching Research Division, Oregon State System of Higher Education, Monmouth

This study was designed to determine the attentive behavior and viewing preferences of preschoolers for national ETV programs designed for preschoolers. The relationship of socioeconomic and ethnic background, age, and sex to viewing behavior was also investigated.

An observational scale was developed to provide reliable measures for visual, motor, and verbal responses. Three observational techniques were compared, two using distraction devices. In one, a latency measure was made as a preschool child restored the visual and audio input of a program after these stimuli had been diminished. A second technique compared the preference of the child for a moving kaleidoscope image versus the ETV program.

Using the kaleidoscope technique, 24 middle class children, half boys and half girls, half 2-3 year olds and half 4-5 year olds, were observed individually viewing a series of five equal length preschool educational TV programs and one cartoon. Twelve 4-5 year old disadvantaged Mexican-American children were observed viewing the same series of programs.

The results clearly indicate a wide fluctuation in viewing patterns during a program. Comparisons of the relationship between program content and viewing behavior relative to socioeconomic class, age, and sex indicated complex functional relationships. General factors related to viewing preference include novelty, fantasy, and viewer participation.

The implications of the study for observational techniques, ETV program development, and evaluation are discussed.

SCHOOL MORALE IN THE METROPOLIS: PUPIL RACE, TEACHER RACE, TEACHER SATISFACTION, AND OTHER CORRELATES, *Reginald L. Jones*, The Ohio State University

Seventeen hundred six (1706) fourth and sixth graders, from 62 classrooms, completed the School Morale Inventory. The inventory comprises an 84 item scale yielding scores on (1) School Plant, (2) Instruction and Instructional Materials, (3) Administration Regulations and Staff, (4) Community Support and Parental Involvement, (5) Relationships with Other Students, (6) Teacher-Student Relationships, (7) General Feelings About Attending School, and (8) Total School Morale. Teachers provided information on their race, their extent of satisfaction with teaching, the number of pupils in their class by race, the average socioeconomic status of the class, their years of total teaching experience, and years of teaching in the school under study. These variables were correlated with mean class score for each subtest of the School Morale Inventory.

The results revealed, at grade 4, a significant negative correlation ($p < .01$) between percentage of minority group pupils enrolled in the class and pupil attitudes toward student-teacher relationships. At grade 6, the proportion of minority children enrolled in a class was negatively related to pupil perceptions of (1) community support and parental involvement ($p < .01$), (2) relationships with other pupils ($p < .01$), (3) teacher-student relationships ($p < .01$), (4) general feelings about school ($p < .001$), and (5) total school morale ($p < .01$).

Total teaching experience was significantly (and positively) related to virtually all subscales at grade 6 and at grade 4. Satisfaction with teaching was significantly (and positively) related (1) to pupil perceptions of relationship with other pupils ($p < .01$), (2) student-teacher relations ($p < .05$), (3) general feeling about school ($p < .01$), and (4) general school morale ($p < .05$).

Implications of the findings for teacher and pupil placement and other aspects of school programming were drawn.

THE FEASIBILITY OF USING TAPED LECTURES TO REPLACE CLASS ATTENDANCE, *John W. Menne, John E. Klingensmith* and *Dennis Nord*, Iowa State University

This study was undertaken to determine whether listening to taped lectures on an individual basis could replace attendance at live lectures.

Students enrolled in an introductory psychology course at Iowa State University in the fall of 1967 were given the option of attending live lectures or of listening to essentially the same lectures on tape. The taped lectures were made the previous quarter, and each student who so chose was given a tape recorder and a complete set of tapes. The study was replicated in the spring of 1968. In the fall 149 chose tape, 211 chose live lecture; in the spring, 141 chose tape, 197 chose live lecture. For each replication comparisons were made between the experimental (taped lecture) group and the control (live lecture) group on the basis of midterm test scores and final grade attained.

Covariance analysis was used in order to equate the two groups on the available predictor variables. In neither replication was there a significant difference in the mean course grade attained by the two groups, but there was a tendency for the experimental group to do slightly better than the control group. At the end of each quarter the experimental group was requested to fill out a questionnaire reflecting their reaction to the tape method. The reaction was generally favorable. While there are difficulties and limitations in such use of taped lectures, the results of this study indicate that some pressure may be taken off instructors and physical facilities by the selected use of taped lectures.

THE USE OF NONSENSE WORDS TO TEST GENERALIZATION ABILITY IN SPELLING, *Carl R. Personke,* The University of Wisconsin

For tests of phonetic generalization, the use of the actual morphemes of a language always presents the question of whether the test is one of morphemic spelling or phonetic generalization. The most suitable answer to this problem seems to lie in the use of nonsense words. In theory, these should test phonic ability independent of knowledge of morphemic spellings. This approach has been used a number of times, and has been demonstrated to be a reliable measure of a general phonic ability. However, it remains to be demonstrated that a specific generalization may be tested by a particular nonsense syllable.

This paper reports the results of a study in which nonsense words purported to test specific phonetic generalizations. To determine whether word environment had any influence on attempts to use a generalization, proportions of accuracy for each attempt were computed in all cases where the generalization was tested more than one time. In each case, the variance of the category proportions was compared with the variance of the individual observations using a chi-square distribution. Significance was tested at the .05 level of confidence.

Of 17 generalizations tested more than once, only 5 proved to have normal distributions. However, scrutiny of the nonsense words used indicated internal inconsistencies within the generalizations being tested. For example, one generalization tested the spelling of all short vowel sounds. Furthermore, cases of obvious interference by the word environment could be detected in some instances. After correcting for these occurrences where possible, seven of the remaining distributions were demonstrably normal. Of the five remaining cases, there were many indications that better choice of nonsense words and more precise definition of generalizations would have rendered these amenable to testing.

The evidence suggests that a test of nonsense syllables will test the use of specific generalizations if the words are carefully composed, tested, and subjected to analysis in the construction of the test.

AN EXPERIMENT TO DETERMINE THE EFFECT OF A TEACHING PROCEDURE, USING STUDENTS' COMPOSITIONS, ON THOUGHT COMPLETENESS IN FOURTH GRADE STUDENTS' SENTENCES,

Billy L. Turney, North Texas State University; *Gwendel D. Mulkey*, Tarrant County Junior College

This study was concerned with the effect of a teaching procedure, which utilized students' written composition as the primary aid for teaching, on the incompleteness and overcompleteness of thought expressed in the writing of sentences by fourth grade students.

The same teacher taught both a control and an experimental written communications fourth grade class. A total of 52 students was involved. The control class teaching method utilized standard textbooks, workbooks, and duplicating materials. The experimental class teaching approach utilized only student written compositions as teaching aids. The goal of both classes was to improve completeness in sentence expression. The experiment lasted for approximately six weeks.

Randomly assigned to groups, all the students (26 in each group) were subjected to post-test stories for measurement of the two sentence classifications. The post-test measurement revealed a significant difference in means at the .001 level in favor of the experimental group in production of fewer incomplete and overcomplete sentences. Thus, the results indicate that the method of using student written compositions as teaching aids will cause children to produce fewer incomplete and overcomplete sentences in their writing.

TEACHER VERBAL FEEDBACK AND CONTENT DEVELOPMENT,

John A. Zahorik, University of Wisconsin-Milwaukee

Teacher verbal feedback is an important teacher behavior employed during the teaching-learning act. The present study sought to examine the occurrence of various types of teacher verbal feedback in relation to type and phase of content development.

The procedure of this investigation was to obtain and analyze transcripts of tape recorded lessons. The subjects were fifteen third- and sixth-grade teachers and their classes of pupils. To determine feedback usage a twenty-five category feedback instrument developed by the experimenter was employed. Data concerning content type and phase were obtained by applying venture criteria, developed by Smith *et. al.*, to the transcripts. Two phases in relation to ventures were examined: medial and terminal.

There were significant differences between total medial feedback and total terminal feedback for all types of ventures. The types of feedback used to terminate ventures always contained praise or repeat elements and frequently both, lesson progression feedback, and often some response development feedback. The types of feedback used to mediate ventures sometimes contained praise or repeat elements but more frequently consisted of only response development solicitation, original solicitation repetition, or lesson progression feedback.

There were some differences among the specific types of ventures in relation to medial and terminal feedback. The most striking differences occurred in medial feedback. Information venture and concept venture medial feedback contained more praise elements and lesson progression feedback. Causal venture and reason venture medial feedback contained fewer praise elements and more solicitation repetition feedback.

Session 14.8 **Division C**

FEEDBACK AND REINFORCEMENT NO. 2

Chairman To Be Announced

THE INFLUENCE OF MODEL PERFORMANCES AND FEEDBACK ON THE LEARNING OF A COMPLEX MOTOR SKILL, *Bruce D. Anderson*, University of Minnesota

Research on imitative learning has indicated that simple motor acts may be learned from the observation of models. Research on motor learning shows that knowledge of results is a powerful variable in pursuit rotor skills. However, recent research on language development has raised questions about the role of imitation in the acquisition of complex behaviors and has suggested the importance of augmenting knowledge of results with corrective feedback. The influence of these variables was explored in the acquisition of a complex athletic skill, the short golf swing.

Sixty Ss were randomly assigned to five treatment groups in which they received five hours of instruction and practice. A factorial design was employed to assess the effects of providing a filmed model performance and three kinds of feedback (kinesthetic, kinesthetic plus knowledge of results, and kinesthetic plus knowledge of results plus corrective comments). At the conclusion of the experiment, Ss were given a performance test and were rated by an experienced instructor.

No evidence was found to indicate that model performances facilitated learning. However, each of the three kinds of feedback produced significant gains and feedback which included corrective comments produced significantly more learning than the other two. In summary, the results lend support to the importance of corrective feedback from an instructor while the learner executes the skill. Practice without this type of instruction appears to lead to systematic errors which do not seem to be reduced by demonstrations but are more effectively controlled by corrective feedback.

The results of the study suggest that extrapolations from research on the acquisition of simple responses to complex skills should be made with caution and further confirms the value of experiments that test extrapolations from laboratory research in complex educational settings.

PERFORMANCE AS A FUNCTION OF FEEDBACK CONDITION, *Madeleine L. Speiss, Richard Thiel*, and *Eleanor Leventhal*, Southwestern Cooperative Educational Laboratory, Albuquerque, New Mexico

This study was designed (1) to assess, under actual classroom conditions, the effects on learning of error correction compared to no correction, and (2) to determine how different types of correction treatments influence learning. Efforts to improve instruction and learning within ordinary classrooms were evaluated for the broader purpose of developing means by which learning could be accelerated for culturally divergent children in primary grades.

First grade children in 23 classrooms from two states were given the task of learning to print vocabulary words from memory based upon precisely developed methods of instruction, even though the task was considered to be too difficult for inclusion in the regular curriculum of these districts. Negro, Spanish-American, and Anglo children were investigated. The hypotheses were that: (1) correction techniques would result in significantly greater performance than non-correction techniques; (2) pupil self-correction would result in significantly greater performance than would teacher correction of pupils; (3) more immediate feedback and correction would result in significantly greater performance than would greater delay of feedback and correction. These hypotheses afford a seven cell design. Data were obtained from diagnostic tests, sub-tests, post-tests, and retention tests. The lessons lasted 15 minutes each for a total of 40 days.

The following results were obtained:

(1) Children in all six correction conditions performed significantly better than those in non-correction conditions ($p \leqq .01$).

(2) No significant differences occurred between immediate correction and the one day delay group, but both treatments were significantly better than the two day delay group ($p \leqq .05$).

(3) In the short run, teacher correction was superior to pupil self-correction ($p \leqq .01$). In the long run, however, pupil self-correction showed greater retention scores.

CONDITIONS FOSTERING THE USE OF INFORMATIVE FEEDBACK BY YOUNG CHILDREN, *Joyce Teager*, University of California, Los Angeles

In a previous study, it was demonstrated that while young children given complete feedback were better able to transfer to new classification tasks than those given no feedback or only told when the correct answer was selected, no differences for treatment were found on the immediate posttest over the program content. It was felt that perhaps these young children had not yet learned to work for the secondary reinforcement of "being right."

For the present study, it was hypothesized that children given feedback accompanied by a primary reinforcer (raisins) would be superior when compared to those given only verbal encouragement or criticism, and that both these groups would be superior to a group given informative feedback alone. Story booklets about Eskimo life were presented. On each page were 3 black-and-white pictures from which children were asked to select one which answered a question. All 21 children were told that the green feedback indicated a correct picture and the red an incorrect one. After three 15-minute training sessions, a test over program content was given. This was followed by a paired-associate task with new words and pictures. Verbal confirmation of the meaning of the green and red color was discontinued, but half of each reinforcement group continued to receive reinforcement.

No significant differences were found among the treatments on the test over program content. In other words, information alone was sufficiently potent to produce the same level of learning as information augmented by either primary or secondary reinforcement. The major difference was that the children given no reinforcement learned the meaning of the color feedback and were able to use this tool to learn a new set of paired-associates with significantly fewer errors (.05 level) than the groups given reinforcement. While considerably below the information-only group, the subgroup which continued receiving primary reinforcement was superior (.05) to its matched subgroup, whereas the praise subgroup was not.

Session 14.9 **Division C**

MATHEMATICS INSTRUCTION NO. 2

Julian R. Brandou, Michigan State University, Chairman

RAPID ACQUISITION OF A BASIC MATHEMATICAL SKILL: AN EXAMPLE, *Lloyd Hutchings* and *Harold Cook*, Syracuse University

This study attempted to make a preliminary assessment of the effectiveness of a new arithmetical algorithm for addition. The paper assumed a negative posture towards the traditional arithmetic computational procedures, primarily because learning these algorithms not only consumes much of teachers' and students' time and energy, but they are frequently the source of anxiety, frustration and boredom. The development of a set of new basic algorithms to facilitate mastery of basic computations is essential. A set of algorithms has been developed which may facilitate the performance of basic computations. This new set of algorithms should:

1. Reduce the number of operations being performed, thereby reducing the amount of material that has to be memorized;
2. Allow for the simple location of errors, making feedback more readily available to student and teacher;
3. Avoid the confusion which may emanate from the alternation of operations in traditional algorithms;
4. Allow children to quickly attain computational success, thereby reducing the failure experience associated with computational deficiency.

Three teachers, each using the new algorithm and the traditional one, taught 46 4th, 5th, and 6th grade children who were randomly assigned to one of two groups. The teachers had one hour of preliminary instruction in which both the new addition algorithm and the traditional procedure were presented to them as "two new experimental procedures." Each teacher was given five detailed 15 minute lessons and four 5 minute practice drills for each method. These were then used daily in teaching addition. The entire treatment lasted nine days.

The results on a power test revealed that the students exposed to the new algorithm performed significantly better ($p < .005$) than those using the traditional algorithm. This test consisted of 20 addition problems varying greatly in level of difficulty. A second test of individual facts did not show a significant difference between groups. These findings suggest that a brief instructional period can generate rapid acquisition of basic addition computational skill.

AN ANALYSIS OF THE STRUCTURAL LEARNING CHARACTERISTICS OF CHILDREN FROM DIFFERING INSTRUCTIONAL BACKGROUND AS MANIFESTED IN THE LEARNING OF CERTAIN MATHEMATICAL GROUPS, *William Emile Lamon*, University of California, Santa Barbara

This study was designed to fulfill three major objectives:

1. Assessment of the degree of difficulty encountered at several levels of intellectual operation by children between the ages of nine and twelve years in the learning of the structure of selected mathematical groups;
2. Identification of the highest level of learning reached by each child, by program and by grade;

3. Exploration of the effects of variables such as the verbal or the non-verbal I.Q., the age, the instructional program, and socioeconomic status upon certain dependent variables.

Three different systems of instruction and two representative communities were selected. In those communities, 54 fourth graders, 54 fifth graders and 45 sixth grade subjects, each representing an equal female sample of Cuisenaire, Traditional and Dienes systems subjects, were selected from 39 different classes in the Province of Quebec, Canada. Three mathematical group structures were developed and prepared through concrete games. The experimental period was 11 weeks in duration with 100 minutes of instruction time per week. During each instructional session each child was led through six distinct and ordered levels of mathematical thinking.

The results of the study were the following:

1. At the fifth grade level only, progression through the hierarchy of difficulty and complexity represented by the levels of learning revealed a significant difference in performance: the Dienes subjects were superior to both the traditional and Cuisenaire children.
2. The Cyclic Groups experienced the most difficulty.
3. Age or socioeconomic status did not significantly affect the performance at any particular grade.
4. The measures of intelligence affected in different degrees either the overall or the per-level-of-learning performance, both by grade and by program, with the non-verbal I.Q. offering the most interesting results.

The interpretation of the quantitative results, combined with classroom observations, seems to indicate that if fourth, fifth, or sixth graders are given sufficient amount of learning time, they can learn and understand the structure of the mathematical group.

THE RELATIONSHIP BETWEEN FREQUENCY OF TESTING, ARITHMETIC ACHIEVEMENT AND INDUCED TEST ANXIETY IN SIXTH GRADE STUDENTS, *Barton B. Proger, Lester Mann, Raymond G. Taylor, Jr.*, Research and Information Services for Education, King of Prussia, Pa.; and *James E. Morrell*, Centennial School District, Bucks County, Pa.

Classroom teachers often observe that frequent testing encourages higher test achievement. However, no experiment verifying this phenomenon at the elementary level is available. The present experiment involved four groups of 20 students each who were put on four different treatments, three involving repeated testing (daily, every other day and weekly), and one involving only daily practice. This regimen was maintained for five weeks. At the end of this period, and two weeks later, all four groups were tested for arithmetic achievement.

Subjects and teachers were selected and assigned randomly. Teachers were also rotated randomly. Three separate analyses of variance were presented. The first two used achievement (immediate and delayed) as the criteria and controlled for methods, sex and previous achievement. The third analysis used scores on an abbreviated test anxiety scale as a criterion and controlled for methods, sex, previous achievement, and trials. Anxiety was measured at the end of each school week.

On both the immediate and delayed achievement posttests, the daily test procedure yielded significantly higher results than any of the remaining three treatments. All interactions were insignificant, which suggested that low achievers do not necessarily benefit more from non-threatening testing situations. With regard to the analysis of test anxiety scores, trial means decreased significantly in a linear fashion from week to week. However, no evidence could be found that different amounts of testing induce different levels of manipulable process test anxiety. The level of significance used throughout the study was .01.

The authors postulate an explanation of the results of this study on the basis of a physiological and affective arousal theory.

Session 14.14 **Division D**

MEASUREMENT OF YOUNG CHILDREN

Alden W. Badal, Oakland Public Schools, Chairman

IMPLICATIONS OF PIAGET'S WORK FOR TESTING YOUNG CHILDREN, *Edward A. Chittenden,* Educational Testing Service

Many investigators of young children's abilities have attempted straightforward standardization of Piaget-type tasks with varying degrees of success. It is argued that there are other, more fundamental implications in Piaget's work for the general problem of testing young children. These implications are discussed and a program for investigating new testing techniques is outlined.

First, Piaget's work makes it possible to build tests on the basis of developmental theory. His research offers a theoretical basis for test construction and validation.

Second, the *méthode clinique* associated with Genevan investigations suggests some useful new approaches for dealing with responses of young subjects. Unlike most ability tests, Piaget's methods (similar to the psychiatric examination) appear to focus on typical rather than maximal performance; his accounts of child thought do not really depict "ceilings" in ability as much as they depict typical ways of thinking. New techniques for assessing young children could take the typical/maximal distinction into account, and might seek "characteristic" responses rather than "best" responses. Piaget has also proposed criteria for distinguishing between types of clinical reactions, ranging from the "liberated" and "spontaneous" convictions to the "answer at random." The possibility of devising reliable procedures for determining response conviction should be examined.

Third, Piaget's analysis of preoperational thought shows that seemingly contradictory or conflicting ideas can be held simultaneously by the young child. Undoubtedly this contributes to the child's unreliability as a subject. It is argued that this prelogical characteristic should be tested in its own right, and that, more generally, the question of unreliability of young subjects could be clarified.

MOVIE PRESENTATION OF PIAGETIAN TASKS: A PROCEDURE FOR THE ASSESSMENT OF CONSERVATION ATTAINMENT, *Kenneth G. O'Bryan*, The Ontario Institute for Studies in Education; *Frederic J. Boersma*, University of Alberta

Factors of verbal and non-verbal cueing, unskilled manipulation of materials, and deficient or misleading instructions were considered to be sources of error in Piagetian experiments performed by untrained examiners. Accordingly, a 16mm movie of key conservation tasks in length, area, and continuous quantity in solid and liquid was produced. The movie, which closely followed Piaget's formulation of the tasks, incorporated a taped commentary and standard conservation questions. A prepared checklist was used for the recording and classification of subject answers.

After initial experimentation on graduate students in developmental psychology, the accepted version of the movie was subjected to field testing as follows.

Eighty subjects, aged between six years three months and seven years nine months, were randomly assigned to four groups, A, B, C, and D. Groups A and B were first tested by a single skilled examiner using Piaget's clinical method. Two weeks later he retested group A with the same procedure, while group B was tested via the movie and taped instructions. In both pre- and post-tests, no significant difference was found between the groups on any of the tasks. Groups C and D were first tested on the movie, and this was subsequently repeated for group C, while group D was tested clinically. Again, no significant differences on any of the tasks were found between the groups. Test-retest reliability was high in all cases ($r > .85$). It was concluded that the movie presentation of Piagetian tasks did not produce significantly different results from those obtained through a clinical presentation.

VALIDATION OF THE HUMAN FIGURE DRAWING TEST AS A MEASURE OF TEST ANXIETY IN ELEMENTARY SCHOOL CHILDREN, *Patricia Lee Engle* and *Joan E. Sieber*, Stanford University

While drawings of a human figure have frequently been used as indicators of an individual's characteristic anxiety state, past research has not always supported this interpretation. Disagreements between studies may be due to differences in populations used, employment of only a single dependent measure, differences in scoring categories used, or lack of a defensible theory for predicting the effects of anxiety on human figure drawing. In this study, these factors were controlled.

Subjects were 133 middle-class 5th- and 6th-graders. The scoring categories were derived from the literature and agreed highly with other scales previously developed. Dependent measures were the test anxiety and defensiveness scales of Sarason's Test Anxiety Scale for Children, IQ, and behavioral measures of stress in a problem-solving situation. A theory based on an analysis of human figure drawing as a cognitive task requiring skills of representation and integration was suggested. It was predicted that a single-summed score for each subject's drawings would correlate with the dependent measures of anxiety, and that, theoretically, two subscales could be developed which would reflect distinct patterns of anxiety. Each would be differently related to the anxiety measures.

Results indicated that the total human-figure drawing score was significantly related ($p < .01$) to self reports of anxiety and defensiveness. Correlations with behavioral measures were generally in the right direction, but were not highly significant. The two subscales of the original scoring instrument were negatively correlated with each other, and correlated at a low but significant level with some anxiety measures.

It is suggested that this instrument can be used as a part of a larger battery of anxiety tests, but that its predictive power does not justify its use as a clinical tool. Ways in which the drawing task may offer some specific evidence of changes in cognitive processes under high anxiety are suggested.

THE USE OF THE GOODENOUGH DRAW-A-MAN TEST AS A PREDICTOR OF ACADEMIC ACHIEVEMENT,

Glenn B. Schroeder and *Katherine A. Bemis,* Southwestern Cooperative Educational Laboratory, Albuquerque

The data analyzed were computed from a study involving the testing of 335 first grade Anglo and Spanish surnamed children to determine the effectiveness of the Goodenough *Draw-A-Man Test* (GDAM) as a predictor of academic achievement with this particular population.

A sample of 18 first grade classrooms in four public schools in the lower socioeconomic area of a Southwestern city was selected for testing with the GDAM, the *Lorge-Thorndike Intelligence Test,* Form A, and the *California Achievement Test,* Form W. Those Ss ($N = 335$) available for final testing constitute the sample.

An ANOVA between Anglo and Spanish Ss' scores on the LT revealed significance at the .01 level. There was no significant difference between the two groups on the GDAM. Both the GDAM and the LT correlated significantly with the CAT total battery ($r = .31$ and $.54$, respectively), thereby supporting the hypothesis. Covariance of the LT and the CAT total battery revealed a significant difference between the two groups ($p < .01$). No significant differences appeared between the two groups when the GDAM was covaried with the CAT total battery.

The mean difference between these two groups of Ss on the GDAM was 2.63, while the mean difference on the LT was 4.39. This seems to indicate that the GDAM tends to bring these two divergent populations closer together as far as IQ scores are concerned. The GDAM predicted the CAT scores for both groups with the same accuracy as the LT. Although IQ scores might be relative to the measure used, it seems evident that the GDAM might well be used to advantage with this population.

Session 15.2 **Division C**

SOCIAL CLASS, RACE, AND ACHIEVEMENT

James E. Hedgbeth, Michigan State University, Chairman

GROWTH IN ACHIEVEMENT CURVES BY STUDENT SOCIOECONOMIC STATUS AND RACIAL-ETHNIC GROUP MEMBERSHIP,

Tetsuo Okada and *Wallace M. Cohen,* U. S. Office of Education

This paper presents growth curves for different racial and socioeconomic groupings of students utilizing data from a national survey of educational opportunity which was sponsored by the U.S. Office of Education. 4,000 schools and 650,000 students were included in the sample.

The purpose of these analyses was to provide evidence for the differences and similarities of growth in Achievement for different socioeconomic and racial-ethnic groups and for the relative roles that class and ethnicity play in these differences.

These curves were obtained by interpolating and interlocking the Verbal Ability, Reading Achievement, and Mathematics Achievement tests administered at grades six, nine, and twelve. Three levels of student socioeconomic status were used and seven racial-ethnic groups: white, black, Mexican-American, Puerto-Rican, Indian-American, Oriental-American, and Other. Breakdowns are also given for four regions of the country and by metropolitan and non-metropolitan areas.

Results show that white and Oriental groups surpass the other groups in their achievement levels, but that the curves are similar for all groups except blacks. For blacks, the Mathematics Achievement curve tapers off or flattens in moving from grades 9 to 12. When the racial and ethnic groups are stratified by socioeconomic status (SES), the achievement levels are higher for SES groups, but the achievement levels of whites and Oriental-Americans still surpass those of comparable SES groups. The curves for the different SES and racial-ethnic groups are roughly comparable, except for Mathematics Achievement where lower SES groups, particularly non-white, show a marked plateau in moving from grades 9 to 12. Results are discussed in terms of differential curriculums pursued by different SES and racial-ethnic groups and differential performance requirements of schools attended by these groups.

SCHOOL AND STUDENT ATTRIBUTES THAT CONTRIBUTE TO STUDENT ACHIEVEMENT AND ASPIRATIONS, *George W. Mayeske* and *Wallace M. Cohen*, U. S. Office of Education

This paper presents the results of analyses designed to develop student background and school factors and to determine their relationship with school achievement and attitudes. The data are taken from a national survey of 4,000 schools and 650,000 students sponsored by the U.S. Office of Education.

About 400 student, teacher, and school variables were reduced into indices using factor analytic techniques. Using the school as the unit of analysis, a set of student body indices (B), comprised of Socioeconomic Status, Family Structure, and Racial-Ethnic Composition, and a comprehensive set of thirty-one school indices (S), comprised of the school's Facilities, the teacher's Training, Experience, Socioeconomic Background, Attitudes Towards the students, Racial-Ethnic Group membership, etc., were related (using multiple regression analysis) to a number of school outcome indices. Included in the indices were the student's Expectations for Academic Performance, Attitude Toward Life, Educational Plans and Desires, Study Habits, and Achievement. Analyses were conducted for grades 1, 3, 6, 9 and 12.

Regression analyses showed that at the higher grade levels all of the outcomes became more predictable, and the variance common to the two sets (B and S) increased for Educational Plans, Attitude Toward Life, and particularly for Achievement. At the twelfth grade 75 percent of the variance in Achievement was bound up in the B and S overlap. Variance common to B and S declined at the higher grade levels for Study Habits.

These results indicate that the influence of the schools is bound up with the kinds of students that they get initially. The schools play an important role in promoting student achievement and aspirations, but the extent of this involvement is greater for higher socioeconomic status and white students than for lower socioeconomic status and non-white students.

GROUPING PRACTICES WITH DISADVANTAGED PRIMARY SCHOOL CHILDREN: THE SECOND YEAR, *Dean K. McIntosh*, University of Hawaii; *G. Phillip Cartwright*, The Pennsylvania State University

The purpose of this project was to determine which, if any, of three administrative grouping plans for disadvantaged children in grades K-3 is most effective in terms of academic achievement and administrative feasibility.

The grouping plans were designed as follows: (A) Heterogeneous, self-contained—children were placed in classes according to chronological age; (B) Homogeneous, self-contained—children were grouped by IQ and achievement; (C) Flexible—children were placed in classes by MA but shifted between and within classrooms for different small group activities.

All the children in the school were randomly assigned to one of the three treatment groups. Groups A and B each contained approximately 120 children and four teachers; group C contained about 145 children and five teachers. Children were considered disadvantaged on the basis of reports of school officials, residence in a federal-state low-income housing project, and mean PPVT IQ scores of 88. (Based on 360 children in school in September 1966.)

First and second year academic evaluations were made. No significant differences among the three groups were found either by total groups or by age groups five through nine. Also, there were no significant differences in attendance.

Evaluation of administrative feasibility was subjective, but suggested that three separate administrative grouping plans in one school is possible. It further suggested that flexible grouping procedures are possible with disadvantaged children of kindergarten age.

RELATIONSHIP OF CLASSROOM GROUPING PRACTICES TO DIFFUSION OF STUDENTS' SOCIOMETRIC CHOICES AND DIFFUSION OF STUDENTS' PERCEPTIONS OF SOCIOMETRIC CHOICES, *Robert P. O'Reilly* and *Gregory J. Illenberg*, New York State Education Department; *Alfred P. MacDonald*, West Virginia University

The objective of the study was to determine whether variations in procedures for forming instructional groups were related to differences in the spread or diffusion of students' sociometric choices and students' perceptions of selection by others.

Ss for the study were 581 students in grades 7 through 12, comprising 20 intact classrooms in four rural schools. Subjects regularly received instruction in one of four classroom grouping procedures: student-pair learning, graded (N = 62); conventional, nongraded (N = 204); conventional, graded (N = 231); and student-pair learning, nongraded (N = 84).

*S*s were administered a six item sociometric choice device which allowed unlimited choice with respect to the dimensions of "liking," "school competence," and "social power." The first three items required selection of others along these three dimensions; the remaining three items asked the *S* to select those whom he thought had selected him. The index of diffusion for the six items was obtained by dividing the number of selections made by a *S* by the number in the class minus 1 and multiplying the resultant score by 100.

The data for the four treatment groups were analyzed using a one-way analysis of variance of the treatment group scores for each item in the sociometric device. The between-groups F-ratio was significant at beyond the .01 level on student perception of "own" social power; the between-groups F-ratios for the other five dependent variables exceeded the .001 level. Inspection of the means for the four treatment groups indicated that greater diffusion of sociometric choices and perception of choices made by others were associated with membership in nongraded and conventional classrooms.

The implications of the findings are discussed in relation to the research relating diffusion of sociometric choice to satisfaction with school and to school achievement.

Session 15.3 **Division C**

CONCEPT LEARNING NO. 2

Walter Ehrenpreis, University of Pennsylvania, Chairman

CONCEPT DISCRIMINATION AS A FUNCTION OF VARYING RELATIONSHIPS BETWEEN THE RELEVANT AND IRRELEVANT CONCEPT, *Russell E. Ames, Jr.*, Indiana University; *Laurence D. Brown*, Indiana University

Previous research on negative instances in concept learning generally has not considered the set of negative instances as themselves forming a concept. In such a case the task involves concept discrimination as well as concept attainment. Given these conditions the irrelevant concept can be either related or unrelated to the relevant concept. This study investigates the differential effects on verbal concept attainment of four kinds of relationships between the relevant and irrelevant concepts, and hypothesizes a main effect for relationships.

A 2 x 2 x 4 design was used; the main dimensions were four relationships (opposite, common class membership, equivalence, and unrelated), plus the two control dimensions of high and low dominance, and graduate and undergraduate subjects. Forty summer graduate students and forty regular undergraduates at Indiana University were administered a card sorting concept attainment task. Each task consisted of eight words taken from the Underwood-Richardson word norms. Time was the dependent measure.

Analysis of variance showed significant main effects at the .05 level on dominance, on groups of students, and on relationships. No significant interaction effects were found. *S*s formed the positive concept in less time under the high dominant condition than under the low dominant condition. Undergraduates completed the task in less time than graduates. Duncan's New Multiple Range Test showed that *S*s formed the positive concept (black) in less time when the negative concept was unrelated (round) to the positive concept than when the set of negative instances was a common class member (green), an equivalence (dark), or an opposite (white) of the set of positive instances.

Concept discrimination is less facilitated when a relationship exists between the relevant and irrelevant concepts.

PROMPTED TRAINING IN CONCEPT IDENTIFICATION, *J. Kent Davis*, University of Victoria

Two experiments were conducted, the primary purpose being to examine the influence of prompted training on concept identification performance. Prompted training involved the presentation of a cue which provided a *S* with information concerning the correct alternative prior to his response. Since this training procedure has been found to facilitate other types of learning, it was hypothesized that it would also facilitate concept identification.

In both experiments the concept identification task required *S*s to classify into four categories patterns which could vary along as many as seven bi-valued dimensions. Each category represented one of four possible combinations of values from two relevant dimensions. Subjects receiving the prompted training were given a prompt on the first trials and thereafter proceeded in a trial-and-error fashion. All *S*s were tested individually and run to a criterion of 16 consecutively correct responses. In Experiment I, 60 college students solved a concept identification problem in a 2 x 3 x 2 factorial design with the following variables: (a) type of training: prompted or trial-and-error, (b) task complexity: 1, 3 or 5 bits of irrelevant information, and (c) problems: 2 problems which varied in terms of the relevant dimensions. In Experiment II, 30 college students solved a concept identification problem in a 3 x 2 factorial design with the following variables: (a) number of prompts: 0, 16 or 32, and (b) problems: 2 problems varied in terms of the relevant dimensions.

The major results of Experiment I were: (a) prompted training facilitated concept identification performance, (b) performance was an increasing linear function of complexity, and (c) the interaction of type of training and complexity was not reliable. The results of Experiment II indicated that both prompted conditions facilitated concept identification, but that there was no significant difference between the 16 and 32 prompt conditions. The facilitating effects of prompted training suggest that prompting aids concept identification by reducing the memory requirements of the task and by providing an optimum amount of time for information processing.

THE EFFECT OF THREAT AND ANXIETY ON CONCEPT FORMATION, *Dean W. Forbes*, Portand, Oregon, Public Schools

This study explored the effect of threat and anxiety on concept formation in an examination setting. Concept formation involving materials at two levels of complexity was studied under four different patterns of threat and visibility of performance report. Twenty-four classes from three different metropolitan high schools with different patterns of socioeconomic class and parental occupation were involved. In each school a complete replication of all eight combinations of treatment and complexity was carried out.

Change in performance between pre and post experimental concept formation tests was related to measures of anxiety and success as well as to experimental treatment.

The following results occurred. For complex materials a significant decrement in concept formation occurred under conditions of threat. Significant interactions appeared between degree of success and threat, and also between success and visibility. Similar results did not appear with simple materials. Girls tended to have higher general and test anxiety scores than did boys but there was no sex difference in performance on the experimental tests. IQ was found to have a low negative correlation with test anxiety but not with general anxiety. Neither test nor general anxiety was related to reading or arithmetic achievement. Decrements in concept formation score were not related to anxiety scale scores but were related to anxiety producing situational circumstances. They were also not related to IQ or school achievement.

The results did not show anxiety to be a facilitating or energizing state and are interpreted as not confirming the drive state theory of anxiety. As an alternative, they do confirm the theory that anxiety is a disruptive state, which is triggered by factors involved in a specific situation, and involves learned behavior patterns.

CONCEPT ATTAINMENT AS A FUNCTION OF INSTRUCTIONS AND A PRINCIPLE, *Dean L. Meinke*, The University of Toledo; *Herbert J. Klausmeier*, The University of Wisconsin

The purpose of this study was to investigate the effects of instructions that convey information, a strategy, and a principle upon the acquisition of concepts that vary in relevant and irrelevant dimensions.

Three types of instructions were used. Minimal instructions conveyed only that information necessary to proceed with the concept attainment problem. In addition to the minimal instructions, structure instructions conveyed information concerning the stimulus materials. Strategy instructions provided a plan for selecting key stimulus cards efficiently. One half of the *S*s were given a principle which was designed to facilitate information utilization. Further, *S*s were assigned at random to four conditions defined by the concept problem to be attained. The four conditions were two attribute problems with four irrelevant dimensions, three attribute problems with three irrelevant dimensions, four attribute problems with two irrelevant dimensions, and five attribute problems with one irrelevant dimension. In each condition *S*s attempted to attain four concept problems.

A 2 X 3 X 4 X 4 factorial design with repeated measures on the last factor was used employing a total of 96 *S*s. Four analyses were completed, one for each of the four dependent variables—time to criterion, number of stimulus cards selected, number of hypotheses offered, and the amount of potential information available to the *S*. Significant effects were found for each of the independent variables. The results of the study demonstrated the efficacy of manipulating instructions to facilitate concept acquisition. The incorporation of a principle facilitated the efficiency of concept acquisition. A significant effect was found for the type of concept to be attained, as defined by number of relevant and irrelevant dimensions.

Session 15.5 **Division C**

LEARNING FROM DISCOURSE: PROACTION AND RETROACTION EFFECTS

Robert C. Craig, Michigan State University, Chairman

PROACTIVE EFFECTS IN MEANINGFUL VERBAL LEARNING AND RETENTION, *David P. Ausubel*, City University of New York; *Mary Stager*, Ontario Institute for Studies in Education; *A. J. H. Gaite*, University of Wisconsin

The principal problem in this study was to ascertain whether the proactive learning of confusably similar material has a facilitating, or at least a neutral, transfer effect on *meaningful* learning and retention. Subsidiary problems included the effect on meaningful learning and retention of overlearning of the proactive material, and the possibility of interaction between the two proactive variables (learning and overlearning) and length of the retention interval (immediate versus 7 days).

The learning material used in this study was a passage on the principle tenets of Zen Buddhism. Prior to learning this material, the two experimental groups studied a confusably similar Buddhism passage once (proactive group) or twice (overlearning group), and a control group studied an irrelevant passage on drug addiction. All groups were given a multiple-choice test of retention on the Zen Buddhism passage both immediately after learning and 7 days later. *Ss* consisted of 143 grade 13 students randomly assigned to the three groups. The data were analyzed by means of a 3 x 2 analysis of variance, with one variable being the treatment received (overlearning, proactive learning, control), and the other variable being the length of the retention interval (Immediate versus 7 days).

Contrary to the typical proactive interference finding in *rote* verbal learning, proactive learning of the Buddhism material neither interfered with nor facilitated meaningful learning and retention of the Zen Buddhism passage. Overlearning of the Buddhism passage also had no significant effect on the meaningful learning and retention of the Zen Buddhism material, and there was no significant interaction between either proactive variable (learning or overlearning) and the length of the retention interval. These findings were interpreted as a stalemate between various specified opposing influences making, respectively, for proactive interference or facilitation.

RETROACTIVE FACILITATION IN MEANINGFUL VERBAL LEARNING: GENERALITY OVER TIME AND EFFECT OF DIFFERING LEVELS OF VERBAL ABILITY, *A. J. H. Gaite*, The University of Wisconsin; *David P. Ausubel*, The City University of New York; *Mary Stager*, The Ontario Institute for Studies in Education

This investigation was designed to explore the generality over time of the phenomenon of retroactive facilitation in meaningful verbal learning, using similar but conflicting materials as learning tasks. In addition, the investigation studied the effect of differing levels of verbal ability on this phenomenon.

The investigation consisted of two experiments. Study 1 used 189 grade 13 students in three experimental and four control groups to assess the effect of varying positions of the interpolated material (a Buddhism passage) relative to the original learning task (a Zen Buddhism passage), and of high and low levels of verbal ability on the phenomenon of retroactive facilitation. The results were analyzed using a multiple regression analysis with verbal ability score as a covariate; planned paired comparisons between means were made. The results showed no significant facilitation effect; high verbal ability subjects scored significantly better on the posttest than those with low verbal ability ($p < .01$), but did not exhibit a significant facilitation effect.

Study 2 used the same materials and 100 grade 13 students in a matched group design in an attempt to resolve the ambiguity between the results of Study 1 and those finding a facilitation effect. A 2 x 2 factorial design was used with subjects ranked by verbal ability and then randomly assigned to matched groups. The results confirmed the findings of Study 1: no facilitation effect was found when a similar but conflicting task was interpolated between an original learning task and a posttest.

These results suggest that earlier findings of a significant facilitation effect occurred by chance or because of inadequate control over level of verbal ability between groups. Thus, it is not yet possible to make recommendations to teachers on how similar and conflicting meaningful verbal materials might best be presented within a lesson or subject.

SIMILARITY AND RETROACTION IN PROSE MATERIAL, *Thomas J. Shuell* and *Walter G. Hapkiewicz*, State University of New York at Buffalo

The results of various studies investigating retroactive inhibition in prose material have been equivocal. While several studies concerned with substantive retention of prose have been unable to show differential retention as a function of similarity, others have obtained facilitation when the interpolated learning (IL) is topically similar to the original learning (OL). These latter studies, however, did not include groups which would eliminate the possibility that this facilitation is the result of positive transfer from IL to the test of OL. The purpose of the present study was to investigate the retention of prose when (1) this possible transfer is taken into consideration, and (2) *Ss* are instructed to compare and contrast the material in the stories.

Two stories on each of two topics were selected. Tenth-grade students read either two similar or two dissimilar stories in a fully counterbalanced design. After reading the first story half of the *Ss* were told to compare and contrast the two stories while reading the second story, while the other half were given neutral instructions. Finally, a 30-item, multiple-choice test on the first story was administered. Three control groups were used. One group had no interpolated learning. One group took the test without having read the story. The third group read one story and then took the test on the story similar to the one they had actually read.

The results indicate that there was no transfer from IL to the test. Those *Ss* reading two dissimilar stories received higher scores on the test than those reading two similar stories, and those *Ss* instructed to compare the stories performed better than those receiving neutral directions. Neither of these differences, however, were statistically significant. The results are discussed in terms of the role of similarity in prose material and the type of retention test (i.e., recognition vs. recall) used.

Session 16.3 **Division B**

ACHIEVEMENT IN PRIMARY GRADES

Bernice Wolfson, University of Wisconsin-Milwaukee, Chairman

AGE OF ENTRANCE INTO THE FIRST GRADE AS RELATED TO RATE OF SCHOLASTIC DEVELOPMENT, *Joseph Ilika,* Lamar State College

This investigation was designed to assess the influence of age of entrance into the first grade on subsequent scholastic rate of development. Developmental theory holds that an early start will not result in significant increases in rates of subject development. This proposition was tested by comparing the rate of development of early and late entrants to the first grade who were matched according to sex, intelligence, social class, and kindergarten attendance. The matching resulted in 41 pairs of boys and 49 pairs of girls who had California Achievement Test Scores for most of the elementary school grades 1-6, inclusively.

The t-test was applied to test the significance of the differences between the mean rates of development of the early and late entrants. The rate of each child's subject development was determined by finding the value m in the solution of the linear equation $y = mx + b$, by using the method of least squares. In this investigation y is the intercept of the slope expressed as reading, spelling, arithmetic, total language, or total achievement age scores, while x is the chronological age, and b is a constant.

The results favored the late entrant boys who had faster mean rates of development than the early entrant boys' reading, spelling, total language, and total achievement age scores. Early entrant boys, however, had faster mean arithmetic age rates of development than the late entrant boys. None of the differences in the previous comparisons of rates of development were statistically significant.

In the case of the girls, the late entrant girls attained faster rates of development in all subjects. The differences between the mean rates of development in reading and total achievement were statistically significant at the .05 and .01 level.

On the whole, the results continue to support the beneficial influence of maturation on scholastic rate of development for children with a mean I.Q. of 101. Further comparisons are necessary to determine whether these findings hold up when the same comparisons are made for high and low I.Q. early and late entrants into grade one.

EFFECTS OF A STRUCTURE-PROCESS APPROACH ON THE DEVELOPMENT OF YOUNG BLACK CHILDREN, *Joe L. Frost,* The University of Texas

The Structure-Process approach to cognition and literacy assumes that certain prerequisite abilities are necessary for concept attainment, and that intelligence develops in a cumulative manner. Consequently, educators should be able to analyze deductively a specific intellectual task into its subordinate *structural units* for instructional purposes. The fundamental *process* of education is the provision of controlled discontinuity operating from a base of continuity. Time, per se, and chronological age, per se, are viewed as irrelevant variables in the *instructional* process, serving only as a backdrop against which events transpire. *Ordinal* bases, versus traditional *normative* bases, then, are critical considerations in programming for young children.

Under the auspices of the Southwest Educational Development Laboratory, the Structure-Process approach was used for curriculum development in two Texas and two Louisiana (rural and urban) early childhood centers enrolling 285 three-, four-, and five-year-old disadvantaged black children during the 1967-1968 school year.

The Teachers Report, an informal checklist of activities based on specific objectives, administered during November, January and May, revealed significant gains in communications, fine arts, and physical activities.

Significantly higher gains were found for girls compared to boys. No significant differences were found on other instruments, suggesting a tendency for teachers to favor girls.

Pre and post administration of the *Preschool Attainment Record* produced significant gains on measures of creativity and physical, social, and intellectual development.

The Slosson Intelligence Test pretest and posttest scores were not significantly different for the total population (disadvantaged children usually follow a cumulative deficiency pattern). Three- and five-year-olds retained initial levels, four-year-olds made significant gains, and six-year-olds lost ground, implying maximum effect of the program for younger children. These results appear to be conservative. Cross-test comparisons and variations in administration of instruments indicate that pretest scores may have been inflated, thus reducing pretest-posttest differences.

PREDICTION OF ACADEMIC ACHIEVEMENT IN FIRST GRADE AFTER PARTICIPATING IN PRESCHOOL PROGRAMS, *Ernest D. Washington,* Rutgers State University

A step-wise multiple regression program utilizing four batteries of tests given over three years was used to predict first grade achievement in three groups of children who had participated in experimental preschool programs. The tests used in the batteries were the Illinois Test of Psycholinguistic Abilities and the Stanford-Binet. The experimental programs were the Bereiter-Engelmann program which focused on teaching academic skills to disadvantaged children, a highly structured preschool under the direction of Merle B. Karnes which focused on the amelioration of learning deficits, and a traditional preschool whose major goal was to foster social skills and capitalize on opportunities for incidental and informal learning.

The independent variables in this study were nine subtests of the ITPA and the Stanford-Binet, while the dependent variables were the total reading, language, arithmetic, and spelling subtests of the California Achievement Tests (primary level). It was found that Stanford-Binet I.Q. and the auditory-vocal-automatic (grammar) were consistent predictors of achievement in reading for the traditional and Karnes preschool programs, while none of the variables were significant predictors of achievement in the Bereiter-Engelmann program. In language, it was found that the auditory-vocal-automatic subtest was a significant predictor of achievement for the traditional and Karnes programs, while visual-motor-association (choosing which of four pictures is related to a stimulus picture) was the best predictor for the Bereiter-Engelmann program. Binet I.Q. was consistently the best predictor of arithmetic performance. None of the independent variables was a consistent predictor of performance for spelling achievement.

A STUDY OF THE KEY VOCABULARY AND THE CLASSROOM INTERACTION VOCABULARY OF DISADVANTAGED CHILDREN ACROSS FOUR REGIONS OF THE UNITED STATES, *Athol B. Packer* and *Joseph J. Shea,* University of Florida

The purpose of this study was to identify and compare the key vocabulary and the actual verbalization vocabulary of young disadvantaged children in several regions of the U. S. A., in order to identify words of high interest and crucial meaning for beginning reading and subject matter curriculum materials.

A total of 20 classrooms with 586 kindergarten and first grade pupils in 4 regions of the U.S.A. were involved. Sylvia Ashton-Warner's approach to developing vocabulary and language art skills was used to elicit a key vocabulary word list from the pupils. A second word list was obtained from tape recordings made of teacher-directed classroom verbal interaction or discussion sessions.

The two lists of words were collected in the manner described and then placed into the following categories: fear, sex, locomotion, and a combination of other areas. These data were then grouped according to the geographic region in which they were collected. Chi square tests for independence between word category and geographic regions were computed for both types of word lists. The two lists of words were also compared category by category, using the test for significant differences between correlated proportions, for the purpose of identifying those categories which were used proportionately more often in actual verbalization as opposed to an expressed intellectual or emotional need. The data yielded many interesting findings with respect to the vocabulary variables.

It appears that the Sylvia Ashton-Warner type of induced key vocabulary can effectively be used in the development of reading curriculum materials for disadvantaged children in the U. S. A.

Session 16.5 **Division B**

CONTRIBUTED PAPERS NO. 2

Chairman To Be Announced

A STUDY OF PLANNING: QUESTIONS TO GUIDE CLASS DISCUSSION AND TESTING BY SECONDARY STUDENT TEACHERS OF SOCIAL STUDIES, *Drew C. Tinsley,* University of Missouri at St. Louis, and *O. L. Davis, Jr.,* University of Texas at Austin

The purpose of the study was to determine the cognitive nature of the questions social studies student teachers plan and compose to guide class discussion and to use as tests in junior and senior high school American History classes.

Sixty-seven social studies student teachers with at least six weeks teaching experience were randomly assigned to one of four experimental groups according to instructional task (discussion or testing) and grade level (8th or 11th). Questions planned by all subjects were based on the same instructional material. Raters coded the questions according to six cognitive processes demanded for dealing with them: memory, convergent thinking, logical thinking, reflective thinking, divergent thinking, and evaluation. Obtained data were treated by ANOVA techniques.

Summary of results:

1. These student teachers planned questions which emphasized memory for junior high school students, and memory and evaluation for senior high school students.
2. Overall, the cognitive objectives of the questions planned did not differ for junior and for senior high school classes.
3. These student teachers did not differentiate the questions they planned for tests and for discussion.
4. The number of questions asked in one category was unrelated to the number asked in other categories.

Implications for program development and research on the significance of questions in teacher education include the following:

1. Specific attention should be directed toward helping teacher candidates develop lesson plans which incorporate questions demanding a variety of cognitive behaviors.
2. Major attention should be given to development of specific understanding of the purposes of questions and skills of classroom questioning in the social studies.
3. Specific attention should be given to development of purposes of testing and to the skills and understanding needed to prepare tests in the social studies.
4. Deliberate attention should be given to help teacher candidates learn how to foster different cognitive objectives in social studies classes.

A COMPARISON OF TEACHER-PUPIL INTERACTION PATTERNS IN E.S.S. (ELEMENTARY SCIENCE STUDY) "DISCOVERY" CLASSES WITH THOSE OF CONVENTIONAL ELEMENTARY SCIENCE CLASSES, *Fred K. Honigman*, NeaRad Corporation, and *Alfred Smeraglio*, Philadelphia Public Schools

Almost without exception, the idea of "student discovery" permeates the new elementary science programs. Although this expression is used with different emphasis and interpretation in each of the programs, it still bears a common thread of meaning for all of the programs. In generalized terms, "student discovery" suggests the establishment of relatively unconventional patterns of teacher-pupil interaction and classroom management. In place of traditional one-way communication (teacher-to-pupil) and one-party initiation of ideas (the teacher's), "student discovery" suggests that much initiation of communication and ideas rests with the students. Moreover, student-to-student communication (previously forbidden) is now regarded as constructive and appropriate classroom behavior. In general, therefore, "student discovery" suggests a radical departure from the traditional teacher-pupil communication patterns in the classroom in the direction of students enjoying unparalleled autonomy and spontaneity of action and thought.

Of all of the new science programs, E.S.S. allows students the greatest freedom of choice and movement and the greatest possibilities for individualized learning activities. E.S.S. thus contains perhaps the greatest potential for meeting the above-mentioned ideals of "student discovery."

To test whether or not teachers and students in E.S.S. classes reflected the unique patterns of communication and behavior suggested by "student discovery," the authors compared 16 classes in which teachers had been trained in the E.S.S. approach with 13 classes of "regular" teachers teaching science, using the *Multidimensional Analysis of Classroom Interaction.* This instrument permits an accurate record to be made of the relative amounts and typical lengths of the various kinds of behaviors performed by teachers and pupils, as well as a record of the frequency of certain important behavior sequences.

Students in the E.S.S. classes demonstrated a significantly greater amount of active participation than their counterparts in conventional science classes, gave a greater number of "high-level" contributions (inferences, opinions, judgments, etc.), performed a larger number of spontaneous (unsolicited) contributions, and performed more student-to-student interactions. E.S.S. teachers asked more questions than their counterparts, selected specific respondents less often, and spent a relatively smaller amount of time lecturing.

ABSTRACT THINKING, A SUBCAPABILITY OF PROBLEM SOLVING IN THE CURRICULUM: THE APPLICATION OF THEORY AND RESEARCH TO CURRICULUM PRACTICE, *Sara W. Lundsteen,* The University of Texas

Gagné's theory of hierarchial curriculum requires that crucial subcapabilities be located, turned into objectives that can be implemented, and tested. The goal of the present study was to investigate the cause-effect relationship between trained and untrained variables of abstract quality in children's thinking as related to a curricular hierarchy designed to improve creative problem solving. The null hypotheses were the following: (a) As measured by 3 tests of abstract thinking, there are no significant differences among scores for a group receiving 23 lessons, once a week, designed to develop the abstract quality of thinking at fifth-grade level and 4 other groups receiving the same amount of instructional time but with 4 differing treatments. (b) There are no significant differences or transfer between the above groups on criterion measures of problem solving, listening, and reading.

A design containing five groups, pretests, and posttests was selected. The subjects were selected from a pool of 45 intact classes whose teachers had volunteered for the experiment, and each subject was randomly assigned to one of the 5 groups. The measurement consisted of a battery of 11 measures including 3 experimental tests of qualitative levels of thinking.

The completed findings tend to reject the null hypotheses and support the research hypotheses of significant differences between treatment groups on all measures. Analyses of variance and covariance and Scheffé test were employed. Differences or rankings favored almost consistently the group trained in qualitative levels of thinking when the difference was not in favor of the treatment group specially trained on the transfer variable in question—listening, problem solving or reading. In the abstract, or qualitative treatment group, differences favored the low I. Q. group on the STEP reading total and on 3 out of 4 subscores ($p < .05$).

In a gross way the Gagné theory of curriculum appeared to be supported with respect to the abstract subcapability. For the children with low I. Q. (62-91), assistance with the abstract thinking quality may enhance reading ability at the fifth-grade level.

APPLICATION OF CONTINGENCY MANAGEMENT TECHNIQUES TO CURRICULUM DEVELOPMENT, *Michael J. Demchik* and *Ernest A. Vargas*, Human Resources Research Institute

There have been few attempts to underpin development of a specific curriculum, especially at the secondary school level, within the framework of a specific psychological theory. Gagné and Duckworth are among the very few who have attempted this at the *elementary* level. The purpose of this paper is to present a secondary high school general chemistry curriculum based on the laws of learning of operant psychology.

Behavioral objectives were first specified at two levels—the information to be learned and the complex behaviors to be developed. Particular teaching modes were designated to expedite obtaining these two types of behaviors. Curriculums were chosen which would best fit a particular mode; for example, the materials for program instruction are not the same as those for divergent teaching. These materials were sequenced with respect to the kind of information necessary for the next level of information and the mode to be used before going on to a different type of behavior. A contingency management system was designed in which extrinsic reinforcers are initially used to build up the learning process; then the material is used as a reinforcer necessary to the next curriculum step. Whenever possible, higher probability behaviors are used to reinforce lower probability ones.

This new curriculum aims to gradually fade dependence on highly structured material toward open-ended study. This curriculum innovation, hopefully, will not only encourage others to take similar steps, but will also rectify some difficulties in teaching harder secondary subjects, thereby reducing the avoidance of these subjects by large numbers of students.

Session 16.6 **Division C**

TEACHER EDUCATION NO. 2

David Lyal Holder, Brigham Young University, Chairman

TRAINING TEACHERS TO DETERMINE LEARNER ACHIEVEMENT OF OBJECTIVES IN READING INSTRUCTION, *Carl Wallen*, University of Oregon; *Jordan Utsey*, State University College at Buffalo

This study investigated the effectiveness of a multi-media instructional program designed to train elementary school teachers to determine learner achievement of objectives in two areas of reading, word recognition and comprehension. The four hour instructional program provides a series of simulated teaching experiences through the use of a self-correctional workbook correlated with video-tape observations of classroom episodes. This is Part I of a two-part program designed to train reading teachers to use a Test-Teach-Test instructional strategy in teaching reading to elementary school children.

Part I is the Test portion of the strategy.

The objective for Part I is the following:
Task. Given an objective in reading word recognition or comprehension stated in the unoperationalized manner commonly found in teacher's guides and manuals, e.g., "phonics—initial consonant b" or "know the meaning of the word stem," write an informal test which could be used to determine if a child has attained the objective.
Criterion. Teachers should write a test having three features: (1) a task requiring a pupil response, (2) the pupil response which will be considered adequate, and (3) a satisfactory resemblance between the task and adequate response suggested, and the categories of word recognition and comprehension defined by the investigators.

The instructional program was administered to 65 student-teachers at the University of Oregon and 82 in-service teachers in the area. A criterion measure was administered before and after instruction.

The instructional program was successful in that few of the participants attained the three criteria before instruction and most attained the three criteria after instruction. The results also indicated that experienced in-service teachers are no more able to determine learner achievement of specified reading objectives than are inexperienced student teachers enrolled in their first methods course.

AN EXPERIMENTAL STUDY OF IN-SERVICE TEACHER TRAINING TO PROMOTE INDUCTIVE TEACHING AND CREATIVE PROBLEM-SOLVING, *Douglas E. Stone,* University of South Florida, Tampa; *Donald E. Mertic* and *Yossel Naiman,* Chicago Board of Education; *Malcolm Provus,* Pittsburgh Public School System

The purpose of this research was to assess the effectiveness of an in-service teacher training program designed to promote inductive teacher behavior. "Inductive" connotes teaching strategies intended to maximize a child's own creative problem solving abilities, permitting him to discover solutions and concepts through his own inquiry.

A sample of 160 volunteer teachers, from 20 Chicago public elementary schools, was selected. One hundred twenty-four of these were involved in the training program, which itself proceeds through a sequence of activities inductively, providing experiences through which teachers may discover for themselves the meaning and virtue of inductive teacher behavior. The training program is the product of several prior years of developmental research, and has been prepared in written sequential form. Thirty-six teachers served as a control group.

Pre and post data were collected on teachers' attitude toward inductive strategies, and on the actual character of their teaching behavior while working with pupils. Observations of teacher behavior were recorded using the Flanders' Interaction Analysis system. Changes were assessed on three dimensions: percentage of "teacher talk," "indirect teacher talk" ratio, and "content" ratio. The statistical significance of observed changes, experimental compared with control, were tested through use of chi square and "t" tests of significance.

Significant changes (.01 and .05 levels), representing increased inductive behavior, were observed for experimental teachers only on all three dimensions of the Flanders'. Significant (.05 level) increase in attitudes favoring inductive alternatives were also observed.

The results indicated the following:

1. Teachers are able themselves to perceive the desirability of changing their own behavior.
2. Fundamental personality changes are apparently not required to effect important changes in teacher behavior.
3. Teachers have a wider potential repertoire of teaching behaviors than they usually exhibit in normal classroom lessons.
4. Local school leadership is an effective change resource.

THE DEVELOPMENT OF A COMPUTER ASSISTED TEACHER TRAINING SYSTEM (PROJECT CATTS): I. PILOT ACTIVITIES AND IMPLICATIONS, *Melvyn I. Semmel*, Center for Research on Language and Language Behavior, University of Michigan

Traditional methods of teacher training are limited by the lack of opportunities for immediate feedback to the teacher in the classroom. Systematic methods for analyzing teacher-pupil interaction in the classroom often require tedious coding, summarizing, analyzing and interpreting procedures. A prototype CATTS is described which is based on a closed loop cybernetic model. The system includes three major components: (1) the *teaching station*, which contains the teacher, pupils, and a CRT visual feedback source; (2) the *observation-coding station*, which includes a behavior coding device and observer-coder; and (3) the *analysis-encoding station*, which currently consists of a PDP-4 computer and hard-copy printout source (teleprinter).

Selected behaviors in the classroom are systematically coded, using one of several coding schemes available (e.g. Bellack, Flanders, Gallagher) by the observer-coder who punches appropriate entries into the behavior coding device ($N \leq 10$ behavior categories). Input is transmitted directly to the computer for immediate summarization, analysis, and feedback of relevant data. Immediate visual feedback is provided to the teacher in the classroom through the CRT in the form of one of several feedback display programs developed (e.g. cumulative percentage of specific behavior across time, "moving window" time line display, deviation from sample time line display). In addition, hard-copy printout of all coded behavior is obtained.

Preliminary demonstrations using the prototype CATTS-I are briefly reviewed. Future modifications in the system are discussed.

Session 16.7 **Division C**

TEACHER CHARACTERISTICS NO. 2

Clark Webb, Brigham Young University, Chairman

THE RELATIONSHIP BETWEEN ETHNIC PREJUDICE AND STUDENT TEACHING BEHAVIOR, *Robert J. Cullen* and *Carl Auria*, Kent State University

This study attempted to determine differences in prejudice between groups of student teachers in early childhood, elementary, and secondary education, how prejudice of student teachers toward Negroes differs from their prejudice toward ethnic groups in general, and the relationship of prejudice to various student teaching behaviors.

A sample of 115 student teachers at a large midwestern state university was studied in the fall of 1967. Ethnic prejudice was assessed with a modified version of Bogardus' *Ethnic Distance Scale* and Hinckley's *Attitude Toward the Negro Scale*. Extensive reports by student teaching supervisors were content analyzed and provided information about each student teacher's behavior. A semantic differential type rating scale was designed and employed to assess particular characteristics of teaching behavior hypothesized to be related to prejudice.

Although no significant differences were found in prejudice scores among the various student teaching groups, it was found as expected that student teachers evidenced greater prejudice toward Negroes than toward ethnic groups in general. Surprisingly, however, the rankings of the mean prejudice scores of the student teaching groups toward 30 ethnic groups were very highly correlated. As hypothesized, prejudice was associated with certain strengths in teaching behavior. For example, willingness to assume responsibility and effectiveness in classroom communication and presentation were negatively correlated with ethnic prejudice. The traits of tolerance, independence, and equalitarianism, as indicated in teaching behavior, were negatively correlated with prejudice for female secondary education student teachers but not for their male counterparts or for the early childhood and elementary education groups.

The findings indicate that ethnic prejudice is reflected in various student teaching behaviors, that the instruments employed have some validity, and that prejudice as a variable in teaching can and should be studied further.

THE CRITICAL THINKING ABILITY OF TEACHERS AND ITS RELATIONSHIP TO THE TEACHERS' CLASSROOM VERBAL BEHAVIOR AND PERCEPTIONS OF TEACHING PURPOSES, *E. Joan Hunt*, Institute for Child Study, University of Maryland; *Marcia S. Germain*, Prince Georges' County School District, Maryland

In this study 39 teachers (kindergarten through twelfth grade) volunteered to take the *Watson-Glaser Critical Thinking Appraisal*. These scores ranged from 59 to 92; the ten teachers with the highest scores (80-92) and the ten teachers with the lowest scores (59-68) were chosen for purposes of comparison. There was no other appreciable difference in the two groups.

Each of these 20 teachers was observed for three one-half hour periods during the teaching day, and all observations were tape recorded and transcribed. Teachers' verbal comments were classified in two ways: according to routine, cognitive memory, convergent thinking, evaluative thinking, divergent thinking, and according to their supportive, nonsupportive, or neutral quality. Questions asked after the third observation were designed to elicit teachers' perceptions of teaching purposes and learning goals.

The t test was the statistical procedure used to test the significance of the mean differences; the level of significance chosen was .05. The high scoring group made a significantly greater number of verbal comments in the areas of convergent thinking, evaluative thinking, and divergent thinking. They also showed larger, but not significantly larger, mean frequencies in the areas of routine and cognitive memory. Thus, it appears that statements made in the routine and cognitive memory categories are not related to a high measured ability in critical thinking; however, statements evidencing the thought processes exemplified in convergent, evaluative, and divergent thinking do appear to be significantly related to a high measured ability in critical thinking.

The high scoring group made a significantly greater number of comments that evidenced support of children and a significantly greater number of neutral comments; however, the latter seems to be related to the significantly greater number of total comments made by the high scoring group.

In answering questions related to the last lesson observed, the high scoring group stated purposes and learning goals that were observable in the lessons taught, stated goals of learning related to academic aims, and stated needs of teacher education that emphasized child development and individual differences to a greater degree than did the low scoring group.

EFFECTS OF COLLEGE AND PUBLIC SCHOOL SUPERVISORS AND STUDENT TEACHERS' BELIEFS, DOGMATISM, AND SATISFACTION WITH STUDENT TEACHING, *Allen P. Hayes*, Auburn University

The purpose was to determine some of the personal influences of college supervisors and cooperating teachers on elementary school student teachers (STs). There were four dependent variables: (1) ST change in level of agreement with John Dewey in the area of philosophic beliefs (measured by pretest to posttest change in Brown's "PBI"); (2) ST change in agreement with Dewey in beliefs about teaching practices (change in "TPI"); (3) ST change in dogmatism (measured by Rokeach's "D Scale"); (4) STs' "Satisfaction with Student Teaching." College supervisors and cooperating teachers each responded once to the TPI, the PBI, and the D Scale. A sample of two hundred twenty STs from four Florida colleges completed these same instruments, both before and after their student teaching experience, and took the Satisfaction Inventory at the end.

Data were analyzed to determine systematic relationships by using multiple linear regression and *U*-tests, and the findings were then subjected to tests of logical coherence.

The single strongest conclusion was that the pattern of beliefs STs held *prior* to student teaching had greater effect on the young women than the external influences of the internship period (such as cooperating teacher, college supervisor, or institutional effects). Nevertheless, numerous systematic external influences on the STs were identified so that it was judged possible, on the basis of the study's findings, to so structure ST-supervisor "matches" that increase in at least two of the four dependent variables probably would result. The multiple regression analysis technique was judged valuable for future research in the area of student teaching. Some ethical problems seem to be posed by this research.

A RATIONALE AND RELATED EXPLORATORY STUDIES FOR VALIDATING THE OHIO STATE INSTRUCTIONAL PREFERENCE SCALE, FORM I, *Sam Leles*, University of Alabama; *C. Kenneth Murray*, West Virginia University; *Harry R. Barker, Jr.*, University of Alabama

The purpose of this paper was to present the rationale for, and the exploratory studies of, an instructional preference scale. The Ohio State

Instructional Preference Scale (OSIPS) was developed to assess two types of learning associated with pre-service teacher candidates: (1) cognitive information about teaching and learning, and (2) attitudes and dispositions which pre-service candidates bring into teacher education programs. These types of learning were sampled in six defined areas:

1. the nature of the learner
2. the nature of content
3. the role of the teacher as a facilitator of learning
4. measurement and evaluation of learners
5. objectives of learning
6. purposes of education

Test-retest reliability studies were computed. Product-moment correlations were .82, .77, and .66. Samples ranged from N = 37 to N = 97.

A predictive validity study using student teachers provided mean differences between OSIPS scores that were significant beyond the .001 level.

A construct validity study (N = 36) using the Teaching Situation Reaction Test (TSRT) resulted in a correlation coefficient of .45, which was significant beyond the .01 level.

Pre-service candidates (N = 132) in methods courses at the Ohio State University (N = 190) and at the University of Alabama (N = 122) provided data which were factor analyzed. A correlation matrix of 50 variables was set up and six factors extracted. Orthogonal rotation was performed with data from seven samples ranging from N = 26 to N = 312. Findings revealed that pre-service candidates:

1. bring information and attitudes to methods courses that are more varied than alike in terms of the six areas represented in OSIPS;
2. have information and attitudes which lack structure;
3. who are males, bring information and attitudes to methods courses that are more similar than different from the information and attitudes of females;
4. bring information and attitudes that are not consistent with accepted principles of learning or with research findings.

Session 16.10 **Division E**

SCHOOL-ENVIRONMENTAL INFLUENCES OF CHILD BEHAVIOR

Chairman To Be Announced

CHANGING TEACHER RESPONSE BEHAVIOR TO THOSE MORE CONSISTENT WITH GOOD MENTAL HEALTH PRACTICES, *Anne M. Eledmann* and *Norma F. Furst*, Temple University

The purpose of this study was to determine whether or not length of school experience affects teachers' ability to deal with difficult classroom situations in ways which are consistent with good mental health practices. Further, this study analyzed the effects on teachers' response behaviors of an experimental course designed to help teachers expand their behavioral repertoire based on making psychological principles operational.

From a compilation of 1500 "difficult" classroom situations gleaned from teachers, a panel of 5 judges categorized the types of difficulties and built a thirty item Questionnaire illustrative of the variety of problems, The Classroom Situation Questionnaire asked teachers to respond to these difficult situations as they normally would in their own classrooms. The judges then classified responses from 173 teachers with varying lengths of experience. The criteria for classification was based on a "diminishing of student, class, or teacher and interruptive of work being done" dimension or the obverse. Subjecting the data to analyses of variances resulted in finding that teachers with varying lengths of service do differ in their responses to difficult classroom situations, with the most productive teacher behaviors occurring at the third and between the sixth and tenth years of service.

Pre and post measures were taken from 156 teachers enrolled in a graduate course designed to help them translate sound psychological principles into appropriate response behavior. Thirty teachers not enrolled in the experimental section were used as a control group. Subjecting the data to analyses of variance of repeated measures resulted in F ratios which indicated that there were significant differences in pre and post measures between the two groups beyond the .01 level, with the experimental group changing in the desired direction and the control group becoming more punitive as time progressed.

RELATION OF CHILD-PARENT EDUCATIONAL EXPECTATIONS AND SCHOLASTIC ACHIEVEMENT TO THE CHILD'S ACTUAL-IDEAL SELF RATINGS AND SELF-CONCEPT: A CANONICAL ANALYSIS, *Dennis W. Spuck* and *Rod Muth,* Claremont Graduate School

This study examined parents' educational and occupational expectations for their children, educational and occupational expectations of the child, and academic achievement scores as they related to two self rating measures and the self-concept of the child. The population of this study consisted of 125 black boys and girls currently participating in a program designed to motivate ghetto high school youngsters to go to college.

Parents' educational and occupational expectations and the children's educational and occupational expectations were all self-reported on a questionnaire. The STEP test was used to measure the students' academic achievement. Actual self and ideal self ratings were established from two passes over an adjective checklist. The adjective checklist contained three variables which were factor analytically derived from a larger list of adjectives; these variables were identified as social, anti-social and academic self ratings. The correlation between the actual and ideal self ratings establishes the measure of self-concept.

A canonical analysis was conducted relating the child's expectation variables to the two sets of self rating measures and then to the self-concept measure. A chi square test of significance, derived from Wilk's lambda, indicated the existence of significant relationships beyond the .01 level. The scholastic achievement and parental expectation variables were then entered into the relationship in a stepwise manner. This procedure enabled the differential effect and contribution of each additional measure to the child's self ratings and the child's self-concept to be established.

SOME EFFECTS OF VARIED EDUCATIONAL PLACEMENT FOR EMOTIONALLY DISTURBED CHILDREN, *Robert V. Turner*, Richmond Public Schools, and *James D. Beaber*, Louisiana State University

The purpose of this study was to compare some effects of placing primary emotionally disturbed children in special education classes separate from secondary emotionally disturbed or brain-injured children.

The study employed a pre-test-post-test control group design. The experimental group consisted of two classes of children (N = 12) who had been diagnosed as emotionally disturbed and programmed together for educational purposes, and two classes of children (N = 14) who had been diagnosed as brain-injured and programmed together for educational purposes. The control group consisted of four classes of children (N = 25) who had been diagnosed as "learning disability," each class containing approximately equal proportions of emotionally disturbed and brain-injured children. These were programmed together or combined for educational purposes. The original sample contained fifty-one children with a chronological age range of 6 through 13.

All groups were taught in special education classes conducted in three public school systems in Virginia. The experimental treatment consisted of separating pupils for instruction by etiology on the basis of clinical diagnosis.

Effects of separate placement were measured by the Slossen Intelligence Test, Gates-MacGinitie Reading Tests, Vineland Social Maturity Scale, A Class Play, Peer Rating and Self Rating, and an arithmetic sub-test of the Stanford Achievement Test. Computerized variance and covariance techniques were employed to control variables found to affect relationships.

Significant differences were found, and implications for intellectual and social functioning related to separate class placement were drawn.

ELEMENTARY SCHOOL EDUCATIONAL ENVIRONMENT: MEASUREMENT OF SELECTED VARIABLES OF ENVIRONMENTAL PRESS, *Robert L. Sinclair*, University of Massachusetts

This study describes the diversity and similarity of educational environment in designated elementary schools. Five environmental variables were selected for investigation. They were termed Practicality, Community, Awareness, Propriety, and Scholarship. Collective perceptions of fifth and sixth grade students toward the selected variables of the environment were used as a source for describing the school atmospheres.

The hypotheses of the study were:

1. There are differences in educational environment among the designated elementary schools when they are measured by the selected variables.
2. There are patterns in educational environment common to the designated elementary schools when they are presented by the selected variables.

The Elementary School Environment Survey (ESES) instrument, consisting of 100 statements about elementary school conditions, processes, and activities, was developed to obtain a measure of student perceptions of the educational environment existing in the sampled schools. The instrument was administered to 2,173 fifth and sixth grade students from seventy-five classrooms in sixteen different elementary schools. Each child received one of four forms that required about twenty minutes to complete.

On the basis of statistical evidence and various descriptions of the environmental variable scores, the two hypotheses of the study were tentatively accepted. The findings of the investigation showed that there were significant differences (.05 to .01 level of confidence) in educational environment among the measured schools. Also, by employing a two-way analysis of variance treatment, significant differences in environment scores at the .01 level were found for six schools selected randomly from the total sample. An analysis of the distribution of the schools on particular statements included in the ESES instrument revealed that schools clearly differentiated on single statements associated with each variable. Profile descriptions for each institution across all variables revealed that the environments among the selected schools were uniquely different on at least one dimension.

Intercorrelations among the variable scores suggested that there were patterns in environment common to the elementary schools. The most distinct relationships were a relatively high positive correlation (.65) between Practicality and Community variables, and a very low positive correlation (.13) between Practicality and Propriety variables. Also, it was possible to identify statements that consistently described environmental characteristics common to elementary schools scoring highest on the variables and elementary schools scoring lowest on the variables. Finally, by examining the clustering of the high and low scoring schools across the variables, seven environmental patterns were identified.

The results of this study, then, suggest that ESES provides a new means to study the school's setting in which children learn and interact. The information provided by this instrument is unique and useful because it equips an elementary school to view itself through the eyes of the students. Further, the data reported about a school's atmosphere should be helpful in the planning of desired changes in educational programs.

Session 16.11 **Division D**

STATISTICAL METHOD AND THEORY

John Milholland, University of Michigan, Chairman

AN ESTIMATE OF THE ACCURACY (OBJECTIVITY) OF NOMINAL CATEGORY CODING, *Allen L. Bernstein*, Michigan-Ohio Regional Educational Laboratory

Many research activities in the behavioral sciences involve coding observed events, documents, or other items into categories according to prearranged protocols. The objectivity of the person doing the coding is usually checked by having another person(s) code the same items, and by comparing the assigned codes for agreements and disagreements. This paper elaborates a theory for estimating the accuracy of individual coders from the percentage(s) of agreement obtained. In the two coder case, the best estimate of the accuracy of observers X and Y, defined as P_x and P_y, is shown to be

$$1.0 \quad P_x = P_y = P = \sqrt{a}$$

where a is defined as the percent of agreement in the codings of X and Y.

When a third coder Z, has categorized the same items, the best estimates of observer accuracy are

$$1.1 \quad P_x = \sqrt{\frac{ab}{c}}$$

$$1.2 \quad P_y = \sqrt{\frac{ac}{b}}$$

$$1.3 \quad P_z = \sqrt{\frac{bc}{a}}$$

where

a = percent of agreement between X and Y
b = percent of agreement between X and Z
c = percent of agreement between Y and Z.

The discussion shows that error factors introduced by the agreement of two people on an incorrect code are minor. A formula is developed to predict the percent of correctness, m, in the three coder case, when correctness on any item is defined as agreement by any two of the three coders, or all three of them. The formula is

$$1.4 \quad m = P_xP_yP_z + P_xP_yQ_z + P_xQ_yP_z + Q_xP_yP_z,$$

where P_x, P_y, and P_z are estimated from 1.1, 1.2 and 1.3, and

$$Q_x = 1—P_x, Q_y = 1—P_y, Q_z = 1—P_z.$$

Empirical evidence is presented to verify that formula 1.4 has predictive power.

The final section elaborates a means of estimating the variance of P_x and P_y in the two coder case and includes a variance table.

AN EMPIRICAL ANALYSIS OF THE EFFECT OF UNEQUAL SAMPLE SIZE ON THE TUKEY STUDENTIZED RANGE TECHNIQUE, *Robert A. Smith,* University of Southern California

The purpose of this study was to analyze empirically the effects of three methods for dealing with unequal group sizes on the Tukey Studentized Range Technique, one of the most commonly employed multiple comparison procedures. When the groups to be compared are not equal in size, three different procedures have been suggested: (1) the harmonic mean of the group sizes, (2) the harmonic mean of the two extreme range group sizes, and (3) the average value of the group sizes.

It was assumed three variables (number of groups in the comparison, magnitude of the difference in group sizes, and the average group size in the comparison) could affect the accuracy of the three recommended procedures. A computer simulation was designed to determine the empirical effects of each of the methods on type I errors under a wide variety of conditions.

The number of groups had no apparent effect on the Tukey procedure with unequal replications, other things being equal. The differences in group size and average group size did affect the accuracy of the three methods to varying degrees.

Each of the methods was robust for some or all combinations of the affecting variables specified in this analysis. The two harmonic mean approximations consistently displayed less variability from nominal significance levels than the average group size method. Within the range of conditions explored in this analysis, the Tukey procedure can be used for group means with unequal number of replications. This verification of its robustness should encourage wider use of the Tukey procedure.

AN EMPIRICAL INVESTIGATION OF THE PERMUTATION t-TEST AS COMPARED TO STUDENT'S t-TEST AND THE MANN-WHITNEY U-TEST, *Larry E. Toothaker,* University of Oklahoma

The purpose of this research was to investigate empirically the probability of a Type I error and the power of the permutation t-test against a one-tailed location shift alternative for normal, uniform, and skewed distributions. Comparisons were made with similar empirical calculations for Student's t-test and the Mann-Whitney U-test.

For each distribution, a computer was utilized to draw 1000 pairs of samples for each of several small sample sizes and, for each pair, to perform the permutations, rankings, or calculations necessary to compute each of the three statistics. Either rejection or acceptance for each of the statistical procedures was recorded. The sampling was done by obtaining one sample from a population with mean μ and a second sample from a population with mean $\mu + \theta$. When $\theta = 0$, the proportion of rejections was the empirical probability of a Type I error. When each of several specified non-null values of θ was used, the proportion of rejections was the empirical power.

The results for the probability of a Type I error showed the permutation t-test to be conservative with respect to the theoretical values and with respect to the empirical values for Student's t-test and the Mann-Whitney U-test for all populations considered. The results for the power showed the permutation t-test to have power values which were generally less than the theoretical values for all populations considered. The power of the permutation t-test was much less than the power of Student's t-test and the Mann-Whitney U-test for the normal and uniform populations, where Student's t-test generally had the highest power values. For the skewed population, the empirical power of the permutation t-test was occasionally greater than the power of Student's t-test and the Mann-Whitney U-test; however, the Mann-Whitney U-test generally had the highest power values.

STUDIES OF HORST'S PROCEDURE FOR BINARY DATA ANALYSIS, *William M. Gray* and *Richard J. Hofmann,* State University of New York at Albany

Most responses to educational and psychological test items may be represented in binary form. There are a large number of situations in which it is of interest to analyze the relationships among binary variables associated with such test items. However, problems such as the following may occur in such analysis, especially when a test purports to measure a unit trait and also is used to discriminate among members of a group: (1) choice of coefficient of interrelationship (Pruzek, 1967) and (2) variations in the number of correct responses across items (difficulty preference of each item) are reflected in the rank of the intercorrelation matrix of the items. That is, the differences in difficulties across items may be represented as extra factors in a factor analysis of the items (Horst, 1965; Pruzek, 1967).

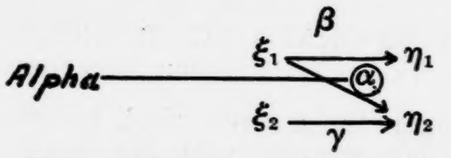
Horst (1965) proposes two general methods of analyzing binary data which will partial out the effects of differences in item preferences: (1) a "least square simplex data matrix solution" and (2) a "least square simplex covariance matrix solution." Of these two, the "least square simplex data matrix solution" was selected for analysis.

Assuming a latent simplex (that part of a system which is due to variation in item difficulty), Horst (1965) applies a least squares transformation to the hypothesized simplex matrix and, in a least squares sense, approximates the binary data matrix and subsequently produces residual data, covariance, and correlation matrices which are "theoretically" free of the effects of the differences in item preferences. An alpha factor analysis of the residual matrices and the original matrices (with a normal varimax transformation) gives results which appear to be at odds with Horst's speculations. Theoretically, after using Horst's procedure, the number of factors extracted from the data should be reduced. However, this does not appear to occur. Also, the item intercorrelations tend to shift toward the negative end of the correlation continuum. Further analysis (canonical, alpha, image, and principal components analysis) and modifications of Horst's procedure(s) are being conducted.

SOME PROBLEMS IN THE APPLICATION OF CROSS-LAGGED PANEL CORRELATIONS, *J. Ward Keesling* and *David E. Wiley*, University of Chicago

In their discussion of Cross-Lagged Panel Correlations, Yee and Gage point out that the technique does not permit unequivocal causal inferences. The reasons they give, however, are not valid. The present authors believe that the difficulty arises from the fact that the causal model is not explicitly stated.

We can diagram a typical cross-lagged panel data problem as follows:



ξ's are time 1 true scores
η's are time 2 true scores

Our problem becomes one of estimating α, β, and γ, which are parameters representing causal influences. Here the question of whether one variable increments or decrements the observed cross-lagged correlation is seen to lack importance. We can estimate the direct causal contribution of one variable to another, and the sign of that contribution is not important.

Noting that there will be measurement errors in any collection of data, we may write the following expressions for observed variables y_i:

$$y_1 = \xi_1 + \epsilon_1 \qquad y_3 = \eta_1 + \epsilon_3 = \beta\xi_1 + \epsilon_3$$
$$y_2 = \xi_2 + \epsilon_2 \qquad y_4 = \eta_2 + \epsilon_4 = \alpha\xi_1 + \gamma\xi_2 + \epsilon_4$$

Our model requires the estimation of ten parameters: α, β, γ, $\sigma^2_{\xi_1}$, $\sigma^2_{\xi_2}$, $\sigma_{\xi_1\xi_2}$, $\sigma^2_{\epsilon_1}$, $\sigma^2_{\epsilon_2}$, $\sigma^2_{\epsilon_3}$, $\sigma^2_{\epsilon_4}$. As our variance-covariance matrix of observed variables has only 10 free statistics, we will be unable to test the fit of our model. Even more discouraging is the fact that β may be estimated in two ways: $\beta = \sigma_{y_3y_4} / \sigma_{y_1y_4}$ and $\beta = \sigma_{y_2y_3} / \sigma_{y_1y_2}$. This means that two of the ten degrees of freedom are given up to estimating β, leaving 8 degrees of freedom to estimate 9 parameters.

There are two possible modifications: (1) Let $\alpha = 0$, which indicates a belief that there is no causal influence. (2) Let $\sigma^2_{\epsilon_2} = \sigma^2_{\epsilon_4}$, indicating that the errors of measurement are stable over time. (Of course $\sigma^2_{\epsilon_1} = \sigma^2_{\epsilon_3}$ would also be possible, alone or simultaneously.) Restrictions 1 or 2 provide for a testable model.

Programs which perform maximum likelihood estimation of the parameters in causal flow models are available, and calculations have been performed in examples similar to those of Yee and Gage.

POSTERIOR PROBABILITIES FOR CONTINGENCY TABLES, *Donald L. Meyer*, Syracuse University

A Bayesian analysis of contingency table data is discussed. Using two theorems proved by Lindley (AMS, 1964), it is shown how posterior probabilities for various ordered alternative hypotheses may be calculated following collection of data. The main result is that, if an indifference prior is assumed for the parameters of a contingency table, the posterior probabilities of contrasts of log-odds is distributed multivariate normally with known variances and covariances. Using tables of the bivariate and trivariate normal distribution, one can calculate posterior probabilities for many contrasts of interests. For larger cases, the use of a Bonferroni inequality may suffice.

DESIGN DETERMINANTS FOR NOMOGRAPHING REGRESSION EQUATIONS, *Lee R. Wolfe*, New York State Education Department

The purpose of this paper is to provide the mathematical, theoretical, and procedural steps for constructing nomographs, with the hope that educational researchers will be attracted to their use. The paper includes the geometric and analytic proofs for nomographing multivariant regression equations and the design determinants for two and three predictor variable equations. The theoretical discussion will make clear the procedure for extending the process to any number of variables. The need for such a paper is clear from the complete lack of literature on the subject.

Researchers produce thousands of regression equations each year, but few are disseminated in a convenient and attractive format to potential consumers. The nomograph, or alignment chart, provides a means for making the solution of a regression equation depend only on a straightedge. The equation can be solved quickly, for a large number of cases, without resorting to a desk calculator, computer, or pencil and paper arithmetic. Nomographs can provide the educator with the same computational convenience that engineering and military scientists have enjoyed in recent years. They are relatively easy to construct and inexpensive to reproduce.

The reader of this paper who is not mathematically inclined may ignore the proofs and proceed directly to the practical examples, and thereby acquire the technique.

MULTIDIMENSIONAL, NONMETRIC MEASUREMENTS OF PERCEPTIONS AND PREFERENCES, *Patrick F. Toole*, Department of Public Instruction, Harrisburg, Pennsylvania

In reviewing traditional scaling methods, Greene, Carmone, and Robinson suggest two limitations: (1) "emphasis on the unidimensionality of the attribute space" and (2) "the strong assumptions underlying the input data when one desires interval or ratio scales." Oppenheim and others previously described these limitations, and Getzels proposed another—the assumption underlying authenticity of subject response. The *purpose* of an unreported feasibility study and a state-wide educational-needs assessment was, among other purposes, to demonstrate the applicability of MAPP (Mathematical Analysis of Perception and Preference) to educational problems and to their description. From this, proposed solutions may be derived.

From more than 200 ESEA Title III funded projects in four participating states, an abstract-stimuli array of diverse project activities was selected. With identifications (geographic) removed, the array was presented to staff from four states (N = 40) and OE (N = 10) to test several hypotheses, among them one concerning congruency of perceptions and preferences relating to ESEA Title III intents. Using these responses as pre-(MAPP) data, a study provided treatment and gathered post-(MAPP) data to test participant (same sample) ability to demonstrate institutional-goal-directed behavior on request if pre-data was divergent from OE jury-judged, desired norms. A third application of MAPP will use photographs depicting various growing-up, formal, and informal educational activities to which various audiences (N = 3000) in Pennsylvania, including children, parents, teachers, administrators and community influentials, will respond.

Feasibility study data analysis is complete. It demonstrates the efficacy of MAPP (heretofore used only in Marketing Science Research) as a technique to describe perceptual/preferential profiles, congruency, and divergency of maps (MAPPs) across different decision-making levels. The second study will be complete *circa* December 1, 1968, and the third, December 31, 1968. Feasibility data analysis (Howard-Harris Clustering Routine) reveals evidence of higher-order constructs, as described by Cartwright and Luce, derived from an analysis of inter-related subjects' response patterns rather than the actual responses themselves. Cluster analysis is verified by OE and state-level occupational and hierarchical responsibilities of the participants involved.

Since MAPP stimuli do not necessarily require verbal response (need-assessment photographs, for example), cultural difference among different subjects attributable to language backgrounds or verbal (or introspective) skills theoretically will not contaminate subjects' responses. Since MAPP, in essence, attempts to elicit how subjects "feel" about certain stimuli in relation to each other, affective-domain measurement implications are obvious.

Session 17.1 **Division B**

EVALUATION IN CURRICULUM DEVELOPMENT

J. Thomas Hastings, University of Illinois, Chairman

EVALUATION OF INSTRUCTIONAL OUTCOMES: THE USE OF UNSTRUCTURED DATA, *Jeremy D. Finn*, State University of New York at Buffalo

The principles of evaluation of instructional outcomes, formally outlined by Ralph Tyler as many as 30 years ago, have had a strong positive effect on education during this period of time. Recently however, changes in thinking about educational practices have made extension of the "traditional" evaluation paradigm imperative. An evaluation paradigm is currently called for which in particular will allow for the valid assessment of a) outcomes of instruction other than the amount of cognitive learning, and b) achievement of special groups of students, such as nursery school pupils or individuals with language deficiencies who may not be amenable to measurement through usual paper-and-pencil procedures.

Such a system of evaluation can be attained through application to the classroom situation of a number of the principles of "assessment," in the sense that the term has been employed by the OSS Assessment Staff in *Assessment of Men.* In particular, both a quantity and variety of types of data are needed by the evaluator in order to make generalizations about the changes in behavior that are taking place during the course of the semester or school year. In order to achieve a more flexible system of evaluation, modifications in the usual testing model are suggested. For one, the testing situation and stimuli items cannot be restricted to standard format. The stimulus situation must be allowed to be naturally occurring, and may even be varied for different students. Second, the response required of the student to any given stimulus must not be severely restricted in form, and may even be "free flowing." Implications of these suggested principles and examples of educational situations in which they may be employed are given.

A set of operational procedures for a method of assessment in the classroom is developed. These include the recording of relevant data and the subsequent projection of the data onto a 'behavior dictionary' in order to obtain valid and reliable measures of the attainment of course objectives. Completed studies which suggest procedures for selecting relevant indicators of achievement and which describe, for example, the collection and projection of language data onto a "behavior dictionary" are related. Finally, an educational application of the extended evaluation system, to be conducted, is briefly outlined.

CONSIDERATIONS GERMANE TO THE IDENTIFICATION OF NON-SPECIFIED FUNCTIONAL OBJECTIVES AND DYSFUNCTIONAL OUTCOMES IN CURRICULUM EVALUATION, *Hulda Grobman*, New York University

In curriculum evaluation, recent emphasis has focused on specification of behavioral outcomes that are immediately observable, are specific to a given subject area, and are positive or desired. These are generally cognitive outcomes, since affective outcomes are less immediate and less readily measured. Little attention has been directed to identification and measurement of outcomes enhancing educational endeavors but not specific to a given curriculum, or to outcomes which impair the achievement of desired ends, where these are not the antithesis of a desired outcome. (The latter may be a direct outcome of a given curriculum treatment, i.e., a Thalidomide effect, or an incidental concomitant.)

No course is neutral in its impact in terms of implementation of many unspecified outcomes. Courses either contribute to or detract from such outcomes. For example, schools assume responsibility for general objectives relevant to political citizenship, individual personality development, development of basic tool skills of communication, etc. While some general skills may be assigned to such specific subject areas as history and government or English, obviously commitment to democratic values cannot be subsumed entirely under the aegis of American history and government, and communication skills are enhanced or discouraged outside as well as during English class. In terms of dysfunctional outcomes, while attention has been given the problem of illustrations in school books as a source of racial prejudice and alienation of minority group children, and historians have been concerned with negative outcomes resulting from unrepresentative treatment of U. S. history, these are only a few of the possible dysfunctional outcomes of curriculums. Evidence of others is legion, but is rarely investigated systematically in evaluation of curriculums.

This paper will explore some areas that lend themselves to study of non-specified functional and dysfunctional outcomes, existing and needed procedures for such investigations, and some difficulties involved in such investigations.

THE MEASUREMENT OF ENTRY SKILLS: AN ESSENTIAL GUIDE IN SETTING INSTRUCTIONAL GOALS, *Rodney W. Skager*, University of California

Among curriculum experts, it is axiomatic that teachers fail to attend to the entry skills of their students. The present research suggests that under certain circumstances teachers may even emphasize skills already attained.

During an evaluation of a curriculum development project, two forms of a forty-item test of approximately thirty mathematics skills were administered at the beginning and end of the school year to random halves of over 700 seventh grade students in three schools. The students were of normal aptitude, but at least one year behind in mathematics achievement, and they took either "traditional" or "experimental" mathematics curriculums in schools representative of an "inner" city metropolitan area.

After posttesting, teachers were asked to judge the likelihood that instruction in their classes would have facilitated students' ability to answer each item. Mean ratings for each item were calculated for each group of teachers (e.g., experimental classes, school X, form A). For each test form the proportion of students passing each item at pretest was correlated with the mean teacher rating. If teachers had concentrated on those skills possessed by few students at pretest, the resulting correlations should have been negative. Instead, *positive* correlations were observed for all twelve groups. Moreover, there were always higher correlations for teachers of experimental classes ($\bar{r} = .52$) than for teachers of comparison classes ($\bar{r} = .23$).

These results, in part confirmed by classroom observations, suggest that in certain situations teachers tend to emphasize familiar, rather than new, content, either because they desire that students experience success, or because it is difficult to motivate learning in marginal students. New procedures are apparently needed for informing teachers of the entry skills of their students. Teachers of the educationally disadvantaged (and perhaps others) may inadvertently contribute to retardation by focusing on skills already available, rather than on weaknesses in need of strengthening.

THE FORMATIVE EVALUATION OF A CHEMISTRY LEARNING UNIT, *Peter W. Airasian*, Boston College

The study was an attempt to determine whether hierarchical structures of tasks inherent in a chemistry learning unit were manifested in student response patterns on a formative evaluation instrument. For the study, Scriven's (1967) definition of formative curriculum evaluation was extended to include the evaluation of student learning and teaching effectiveness.

A unit of learning selected from a chemistry text was analyzed by two curriculum specialists to determine the new content introduced in the unit and the behavior level at which the content was to be learned. The *Taxonomy of Educational Objectives—Cognitive Domain* (Bloom, et al, 1956) defined the behavioral categories. Having characterized the tasks to be learned in terms of a content and behavior component, relationships between the tasks at different behavior levels were posited. These relationships attempted to determine lower level tasks which were necessary but not sufficient prerequisites for mastery of higher level tasks. Thus it was hypothesized that students who failed a lower level task would fail all related higher level tasks.

Maximum likelihood procedures were utilized to derive estimates of probabilities of various response patterns given the dependency relationships between tasks at different behavior levels. Likelihood ratio tests, testing the fit of the estimated probabilities to the observed probabilities, were employed.

An item was constructed to evaluate each task in the learning unit. Two chemistry teachers, teaching about 130 students, administered the formative evaluation instruments.

Analyses of student response patterns showed that the percentage of students manifesting response patterns which conformed to patterns dictated by the dependency relationships between tasks consistently exceeded 80 percent. This finding was generalizable across students of different sex and prior achievement, and across teachers who employed different sequences and emphasis in their instruction. Likelihood ratio tests showed that the observed response probabilities fitted the estimated probabilities in nearly all instances.

The hypothesis regarding hierarchies of tasks to be learned was supported. Formative evaluation instruments incorporating hierarchical structures of tasks can pinpoint, by analysis of response patterns, areas of strength and weakness in the curriculum, learning, and teaching.

DIFFERENCES IN GROUP AND INDIVIDUALLY ADMINISTERED TESTS OF THE SAME BEHAVIOR, ***Dr. Henry H. Walbesser,*** **University of Maryland, and** ***Heather L. Carter,*** **American Association for the Advancement of Science**

The experimental science program, Science—A Process Approach, has endeavored to develop a sequence of instructional materials based upon cumulative learning hierarchies. The theoretical model appeals directly to the work of Gagné. Field tryout has been carried out in fourteen trial centers, beginning in 1963 and continuing until the present.

Competency measures were used for the collection of performance data. Each item on a competency measure is associated with one of the specified behavioral descriptions provided as objectives. Approximately one hundred behavioral objectives from the learning hierarchies appear in the instructional materials for any year. Nine classes of performance were operationally defined and are named by nine action words. Each statement of an objective includes one of the defined action verbs which specify the performance class of the behavior to be acquired. Two forms of the competency measures were prepared for the last three years of the program, and it is with these that this paper is concerned. The group measures make accommodations in the task or response mode requested of the subjects so that the instruments can be group administered. The individual measure does not have these task and response limitations imposed by a written format. The question is the magnitude and direction of the differences one obtains in employing the task and response modifications required in the group format.

The findings report the observed differences in the estimates of acquisition for the tryout sample separated for each of the nine performance classes.

Observed differences are also reported for three socioeconomic categories defined in a previously reported investigation (Walbesser and Carter). The group measures were found to depress the performance results in the describing, constructing, and demonstrating classes. The effect was found to be even more exaggerated among the low socioeconomic group for the describing class of behaviors.

Session 17.2 **Division C**

MOTIVATION

Howard Kight, **State University of New York at Buffalo, Chairman**

MEASUREMENT OF MOTIVATION TO ACHIEVE IN PRESCHOOL CHILDREN, ***Bonnie L. Ballif,*** **Fordham University, and** ***Dorothy C. Adkins,*** **University of Hawaii**

The purpose of this research was to develop an instrument that would measure motivation to achieve in preschool children. Based on extensive investigations of testing techniques and formats, such a measure, *Gumpgookies,* was designed. This instrument is presented in story format, and focuses on the behavior of imaginary figures called Gumpgookies. Each child is told that he has his own Gumpgookie and is asked to identify it. His Gumpgookie is identical in appearance to other Gumpgookies, but behaves only as he behaves. Each item describes two Gumpgookies responding differently to a semi-structured situation, providing dichotomous options that indicate the strength of learned responses hypothesized to be the constituents of motivation to achieve.

In the initial trial of the test, 200 such items were administered to 182 children, 114 from Head Start and 68 from private preschools. Item analysis assisted in the selection of the most promising 120 items, based on both the total sample and a subsample of Head Start children identified as highly and lowly motivated to achieve.

Factor analysis techniques eliminated an additional 20 items, provided some evidence for the seven-factor structure hypothetically underlying the responses, and tentatively identified three second-order factors.

A revised *Gumpgookies*, consisting of 100 items, was then administered to 330 four- and five-year-old children, 160 from Head Start and 170 from private preschools. The stability of the individual item indices, as well as the reliability of the entire composite, was thus cross-checked, and the item difficulty indices and discrimination indices were determined. The new estimated reliability (Kuder-Richardson Formula 20) was .88.

A number of factor analysis techniques were explored, and they provided additional clarification of the underlying factor structure.

This instrument is necessary in order to determine classroom variables related to motivation to achieve in preschool children.

ACHIEVEMENT MOTIVATION OF SEGREGATED BLACKS, DESEGREGATED BLACKS, AND WHITES, *Berj Harootunian*, Syracuse University

After two decades of research on need for achievement, only two empirical studies have been reported that reveal black children to have lower need for achievement than white children. In this study, the achievement motivation of black and white ninth-grade students was compared. There were three comparison groups: segregated blacks, desegregated blacks, and whites, numbering 88, 57, and 624 respectively.

The theoretical formulations of Atkinson (1966) were used to look at the differences in achievement motivation of the three groups.

All students were administered the French Test of Insight, a projective test eliciting need achievement, and the Test Anxiety Questionnaire, devised by Mandler and Cowen. The latter is a self report instrument that measures the construct fear-of-failure. A 3 x 2 analysis of variance (three types of students by two sexes) was carried out on the two sets of data.

For need achievement, significant differences were found for student type ($p < .01$) and for sex ($p < .05$). The segregated blacks had the lowest achievement motivation, the desegregated blacks were in the middle, and the whites had the highest. Girls had higher need achievement than boys. Thus, the segregated black males were the students with the lowest need achievement.

On the Test Anxiety Questionnaire, there were highly significant ($p < .001$) sex differences, with girls reporting more anxiety than boys. The differences for student type were of borderline significance ($.10 < p < .05$), with the segregated black students revealing the highest anxiety or fear-of-failure followed by the desegregated blacks and whites.

The results from both sets of scores are consistent with Atkinson's theory of achievement motivation, particularly in the case of the males. These findings and their implications are discussed in terms of motivation theory and school practice.

SUBJECTIVE RESPONSE UNCERTAINTY AS A FUNCTION OF STIMULUS-TASK INTERACTION, *Gavriel Salomon*, Indiana University; *Joan E. Sieber*, Stanford University

Research has shown that subjective response uncertainty (H*) is a linear function of objective uncertainty (H) up to a certain limit. H* is conceived of as a state of mind in which one experiences a conflict between competing response tendencies aroused by a complex, ambiguous, or surprising (high H) stimulus. H* is defined as a function of the number of competing responses an individual considers, the similarity of their strength, and their total strength. H* often results in information search, the aim of which is to change the probabilities attached to the response tendencies.

In the present study, we tested the hypothesis that H* is a function of the *interaction between stimulus H and the task to be performed by Ss.* The same stimulus ought to produce different degrees of H* under different task requirements. The rationale was that different tasks direct attention to different stimulus components, some of which may not carry surprise or ambiguity. Under certain tasks even extremely ambiguous or unstructured stimuli may not arouse H* when the potentially H*-arousing elements are not relevant.

One hundred and twenty subjects were shown either a structured, logically sequenced film (SF) or an identical but nonstructured film, spliced at random (NSF). Two tasks were required: to record as many cues as possible (CA) or to generate as many hypotheses as possible about the story of the film (HG). The same design, after rotation, was repeated with a second film.

The results showed a clear disordinal interaction between structure of stimulus and task to be performed as measured by the average H*. The dependent variable—average H*—was computed on the basis of the response distribution obtained in each cell (N in each cell = 30). CA with the SF produced significantly higher average H* than CA with NSF, while HG with NSF produced significantly higher average H* than HG with SF.

The conclusion was that to obtain high H*, and consequently curiosity, under cue-attendance requirements, structured learning material is needed. When hypothesis generation is required, nonstructured learning material is more appropriate.

EFFECTS OF ACHIEVEMENT MOTIVATION AND TEST ANXIETY ON PERFORMANCE IN PROGRAMMED INSTRUCTION, *Kenneth Shrable* and *J. M. Sassenrath*, University of California, Davis

The theory of achievement motivation projects an image of an achievement-oriented personality and of a failure-threatened personality. The achievement-oriented person tends to choose tasks in the moderate range of difficulty. The failure-threatened personality, when given a choice, will defend himself by undertaking easy or difficult activities. The present study will attempt to relate strength of achievement motivation and test anxiety performance on a programmed learning task in which the difficulty of all items is kept low.

114 upperclassmen were administered an achievement motivation, a test anxiety, and an initial achievement test, and were categorized as being either high (above the median) or low on each. This procedure defines the experimental variables. *S*s then completed a programmed instruction task and their performance on three dependent variables was assessed:

(a) time, (b) errors, and (c) retention. The results theoretically most interesting are the following: (a) there were no significant differences in errors due to achievement motivation; (b) high test anxious *S*s have fewer errors than low test anxious *S*s; (c) the interaction effect of achievement motivation and test anxiety on errors was significant; (d) high achievement motivated *S*s have lower retention scores than low achievement motivated *S*s; and (e) there was no reliable difference in retention due to test anxiety.

These findings raised certain problems of interest in using programmed instruction. Many programs are written so that most students are able to succeed on each frame. However, the programmed materials must maintain student interest, i.e., *keep the student motivated to perform.* The present study indicates that when item difficulty is made very easy, motivation to perform is increased for avoidance *S*s and decreased for achievement-oriented persons. Recent modification of the theory of achievement motivation suggests that need for different programs for achievement-oriented and failure-threatened *S*s. Theoretically, following success the decrement in avoidance will be greater than the decrement in achievement motivation so that high avoidance *S*s will be encouraged by success. On the other hand, achievement-oriented persons will tend to relax after success. To maximize motivation, this latter group should experience failure on approximately one-half of the items.

Session 17.3 **Division A**

POLITICS OF EDUCATION AT THE STATE AND NATIONAL LEVEL

Jay D. Scribner, Harvard University, Chairman

STATE POWER AND LOCAL DECISION-MAKING IN EDUCATION,

David L. Colton, Washington University (St. Louis)

The purpose of this study was to develop empirically-based concepts suitable to the analysis of state influence on decision-making in local schools. A substantive case methodology was employed. Data were collected from nineteen schools which contracted with the state of Illinois to operate "demonstration centers." Ten state policies were identified, and local compliance with these policies was then measured. Where compliance was found, interview data and documents were searched for statements indicating state influence on local decision-making. The statements were then subjected to content analysis.

A conceptual framework developed by French and Raven was found to fit the data. Local decisions to comply with state policy were based on local perceptions that the state held one or more of the following types of power: (1) reward power—the state's ability to provide money, praise, know-how, and other locally-valued services; (2) coercive power—the state's ability to withdraw or withhold valued support; (3) legitimate power—the state's ability to evoke norms of deference; (4) referent power—the state's ability to inspire local feelings of identification with state policy; and (5) expert power—the provision of ideas and information to local decision-makers. A major conclusion of the study is that state control can be viewed as a dyadic relationship involving five types of power resources at the state level and perceptions of these resources at the local level.

Utilization of the five types of state power appears to be affected by the following conditions: (1) developmental factors within the state program, (2) the substantive character of the decision to be induced locally, (3) personality factors among state personnel, (4) the availability of power resources, (5) time and space constraints (e.g. opportunities for face-to-face contacts), (6) the organizational personality or climate of individual schools, and (7) feedback to the state.

The study has practical significance for individuals concerned with strengthening state education agencies. The study also has implications for the development of a theory of intergovernmental relations in education.

SELECTION PROCEDURES AND THE CHARACTERISTICS AND ROLE EXPECTATIONS OF STATE SCHOOL BOARD MEMBERS, *Gerald E. Sroufe*, Claremont Graduate School

The purpose of this study was to test empirically some of the numerous "propositions" normatively posited regarding the consequences of various methods of selecting state school board members, and to investigate the extent to which board member role-expectations might be associated with different methods of selection.

For purposes of formal examination, selection procedures were considered the independent variable, characteristics and role expectations the dependent variables. A four-fold typology was developed to assist in classifying the variety of selection procedures used in the fifty states. Seventy per cent of the total population of state school board members responded to a mailed questionnaire consisting of three parts: 1) Who Are The Board Members (i.e., biographical characteristics), 2) Becoming A Board Member (i.e., analysis of the process of appointments and elections), and 3) The Role Of The Board (i.e., expectations of the board regarding their role in the state educational system). Statistical analysis (.05 decision rule) revealed the mythological quality of our presumptions regarding the consequences of various selection procedures. For example:

1) There is little absolute and no statistical difference in characteristics such as income, vocation, education, or organizational memberships.
2) Appointed board members show more propensity to leave the board for partisan political office than elected members.
3) "Partisan" campaigns are well-characterized as non-events; if the "best men won't run" it is not because of the expense, effort, or attention demanded in securing an elected position.
4) Role expectations of members varied little according to the selection variable, but revealed interesting common syndromes and common ambiguities.

The findings suggest that reformers concern themselves not so much with mechanics of selecting state school board members—board members are of a single piece—as with means of providing for appropriate socialization opportunities for them. The position of state board members in the educational system is one of manifest ambiguity, yet they have no opportunity for anticipatory socialization, receive the most inadequate instruction regarding their "role," and receive no support for assertive behavior. Those seeking a larger voice for the state board may more profitably consider this category of problems, rather than selection procedures.

COHESION AND FRAGMENTATION IN HIGHER AND ELEMENTARY-SECONDARY EDUCATION: THE AMERICAN COUNCIL ON EDUCATION AND THE NATIONAL EDUCATION ASSOCIATION COMPARED, *Harland Bloland*, New York University

As the national peak associations of higher and lower education, respectively, the ACE and the NEA represent two quite different ways of promoting national coherence in a decentralized educational system. The ACE, on the one hand, has attempted to coordinate and represent a great diversity of autonomous higher education associations by providing a *federative context* (Warren, 1967) for educational decision making. Its constituent organizational members retain individual goal orientations but cooperate in the pursuit of certain common educational objectives, with the ACE providing leadership and some formal organizational machinery for the coordination of their cooperative effort.

The NEA, on the other hand, has attempted to coordinate and speak for diverse elementary and secondary education groups as members of a single organizational structure. The NEA provides a *unitary context* for educational decision making, within which organizational subunits are "structured in a division of labor for the achievement of the inclusive [organizational] goals" (Warren). While divisions within the organization retain their own subgoals, it is expected that their pursuit of these objectives will be consistent with a primary commitment to the inclusive goals of the larger organization.

In recent years, the "unitary" goal orientation of NEA has resulted in organizational strain, with such subunits as the American Educational Research Association and the American Association for Higher Education seeking new autonomy to pursue independent educational missions. The more loosely organized ACE, on the other hand, has strengthened its coordinative function through the development of new formuli for concerting effort among autonomous higher education associations.

Based on research conducted in Washington, D.C. in 1966, this paper will attempt to describe and account for changes occurring in these two peak educational associations.

A PROGRESS REPORT ON RESEARCH ON "THE DEVELOPING RELATIONSHIPS BETWEEN ELEMENTARY-SECONDARY AND HIGHER EDUCATION IN AMERICAN STATES," *Michael D. Usdan*, Teachers College, Columbia University; *David W. Minar*, University of Washington; *Emmanuel Hurwitz*, University of Illinois at Chicago Circle

This project was designed to explore the developing patterns of relationship between elementary-secondary and higher education in American states. It is not a comprehensive study, but rather a preliminary examination in twelve major states of the major dimensions of the subject, intended to detect problems, sources of tension, and modes of reaction as these are being worked out in practice.

Traditionally, Americans have tended to think of higher and elementary-secondary education as two separate and nearly unrelated spheres, despite the arbitrariness of the dividing line between them. More recently, various pressures have begun to bring the two closer together, sometimes in common effort and sometimes in conflict. Among the major factors working in this direction are generalized concern with the effectiveness of American education, increasing costs, greatly expanded demands for higher education, and growing consciousness of certain kinds of societal needs. The basic questions we are asking are these: 1) As these pressures grow, what forms do they take, i.e., how are the relationships between elementary-secondary and higher education felt? 2) How do social, political, and organizational circumstances influence the particular shape and definition of these relationships? 3) How are these relationships being handled in various states?

Based on our findings, we are prepared to suggest the following:

1. Conflict between elementary-secondary and higher education is increasing, though much remains latent.
2. Most of the strain between levels arises out of three major issues: distribution of fiscal resources, the control of public education at 13th and 14th grade levels, and the character of vocational-technical education.
3. The form and intensity of the relationship between elementary-secondary and higher education is related to structural properties of the state system; relevant aspects include the structure of control for elementary-secondary education, coordinative devices in higher education, and the distribution of functions in the legislative and executive branches.
4. Relationships between levels are influenced by the condition of a state's private interest groups in education.
5. Conflict or suspicion between levels tends (a) to induce more detailed legislative policy-making, and (b) to fragment educational planning.

Session 17.4 **Division A**

CONTRIBUTED PAPERS

Frank Lutz, Pennsylvania State University, Chairman

AUTHORITY OF POSITION AND AUTHORITY OF KNOWLEDGE: FACTORS INFLUENCING TEACHER DECISIONS, *Delbert K. Clear*, University of Wisconsin, Milwaukee

The special characteristics of public secondary school teaching make the professional model of occupational life a useful framework from within which to study the authority relations among teachers and supervisors. The purpose of the research was to determine empirically the extent to which the theories concerning the influence of authority of position and authority of knowledge on professionals could be used to predict teacher behavior. Principals were assumed to be authority of position figures, while Department Chairmen were assumed to be as close to authority of knowledge figures as could commonly be found in secondary schools. It was further assumed that teachers' level of professional orientation would determine their relations with these authority figures.

Specifically, the following hypotheses were made: (1) in instructional matters, a specific subject-matter expert's influence on teachers of that subject will be greater than the principal's; and 2) teachers' professional orientation will be universely related to influence from the principal and positively related to influence from the department chairman.

An experiment was designed which provided overt behavioral manifestations of influence. The data were analyzed and tested for statistical significance, using homogeneity of regression and analysis of variance techniques.

The findings were equivocal. Neither hypothesis was supported by the behavioral measures, which indicated no differences between the two authority figures. However, perception data obtained from the participants suggested that the principal is not a source of influence on instructional matters, while colleagues, not including the department chairman, are.

The experiment presents interesting implications for further research in finding effective sources of instructional leadership and, particularly, in assessing the potential of the principal as an instructional leader. There were very interesting methodological implications from the study inasmuch as it was a true experiment conducted in the field. Certain difficulties in measuring professionalism with currently available scales were encountered and solved.

A FACTOR ANALYTIC INTERPRETATION OF THE ISSUES IN THE FLORIDA TEACHER WALKOUT, *Linda B. Stebbins*, University of Florida; *Marjorie Ragosta*, University of Florida

This study was designed during the statewide Florida teacher walkout to investigate the positions taken on the issues as distinguishing characteristics identifying three of the participating groups: the teachers who resigned (N = 421), the teachers who remained in the classrooms (N = 132), and the substitutes who replaced the resigned teachers (N = 51).

A questionnaire containing a five-point rating scale, ranging from very unimportant to very important, requested opinions on the relative centrality of 18 highly publicized issues proposed by factions concerned with the teacher walkout. The questionnaire was administered to the Alachua County sample representing the Gainesville city schools and the other Alachua county schools in all of the affected levels of education, kindergarten through 12th grade.

A factor analysis of the 18 "key" issues for each of the three groups sampled revealed that: (1) there was considerable overlap among the three groups on three factors: sources of powerful outside influences on education, concern for the quality of instruction, and the school climate as experienced by the teachers; (2) there was more factor structure agreement between the resigned teachers and the teachers remaining in the classrooms than between any other two groups; and (3) there were two issues identified by both the resigned teachers and the teachers remaining in the classrooms as unimportant issues in the walkout: increasing teachers' salaries and opening the possibility for unionization of the teachers.

THE INFLUENCE OF RACE, SEX, AND URBAN VS. NON-URBAN TEACHING SITUATION UPON TEACHER ATTITUDES TOWARD BECOMING ADMINISTRATORS, *Frank P. Besag,* State University of New York at Buffalo

This study attempts to determine whether or not there is any relationship between the above mentioned factors and teachers' attitudes toward their own potential role as administrators.

A list of all of the teachers in a large metropolitan school district was obtained from the district. With the districts' cooperation this list was broken down into eight groups: Male Negro Urban ("urban" is defined as an inner city teaching situation; non-urban is the opposite); Male Negro Non-Urban; Female Negro Urban; Female Negro Non-Urban; and the four Caucasian counterparts. A random sample was drawn from within each of these eight groups. The names and addresses of the teachers drawn was obtained, and an instrument was sent to each (approximately 85% return).

The instrument is made up of two parts. The first part contains three sections: questions relating to demographic data, questions relating to the teacher's educational qualifications, and questions relating to the extent to which the teacher has thought of entering the administrative field and the extent to which that aim has been accomplished.

The second part of the instrument is a scale of alienation developed by the present author. The scale contains seven clusters of items involving attitudes toward own school experience, the future, the society and family, negativism, religion, isolation, and powerlessness.

It was found that there are significant differences between each of the major groups (Negro vs. Caucasians, Male vs. Female, Urban vs. Non-Urban) in demographic factors, desire to become administrators, and attitudes toward their own school experience, as well as on the general alienation scale.

This is even more true when the eight sub-groups are compared. For example, Male Negro Urban teachers have a higher educational level, more teaching experience, and have more often thought of leaving the classroom teaching situation. The exact opposite is true of the Male Negro Non-Urban teacher. Another example, Female Negro Non-Urban teachers by far showed the most desire to become administrators, while Male Caucasian Non-Urban teachers least often had this desire.

SIMULATION IN THE TRAINING OF R & D PROJECT MANAGERS, *Duane H. Dillman,* The Ohio State University; *Desmond L. Cook,* The Ohio State University

The general purpose of the study was to apply simulation to the training of managers of educational research and developmental projects in order to improve their decision-making skills. More specific purposes were to develop principles or guidelines for the preparation of simulation exercises and to develop one or more simulation exercises for use in training programs and project managers.

A literature search was made of applicable simulation areas, including industrial engineering, business management, social and organizational psychology, political science, and education. Construction of the initial simulation exercise evolved over a 16 month period. The exercise required the Ss to use role-playing in an educational organization within a sociopolitical environment. They were initially required to develop a project plan using a systems approach under constraints imposed by a request for a "proposal" from a funding agency. Performance standards had to be determined by Ss as the result of imposed time and cost dimensions.

Use of the exercise in six training programs in the university and training center settings provided feedback for its evaluation. The 177 Ss were asked to evaluate the exercise and its use. They identified the most critical problem as that of appropriate and timely feedback to the five-man simulation teams.

The results suggested that simulation may indeed be a valuable supplementary method for the training of managers of such R & D projects. Principles for the development of such exercises are given along with a discussion of the inherent problems.

Session 17.5 **Division C**

SOCIAL CLASS AND LANGUAGE

Nettie Bartel, Indiana University, Chairman

A LONGITUDINAL COMPARISON OF THE AUDITORY MEMORY CAPACITY OF CHILDREN FROM DIFFERENT SOCIOECONOMIC BACKGROUNDS, *Loren S. Barritt,* University of Michigan

Earlier cross sectional studies (e.g., Barritt, 1967) have found that the differences between the language performance of children from middle class and lower class homes *diminish* as the children grow older. This is contrary to the "cumulative deficit" hypothesis which has been advanced to describe the pattern of differences between successful and unsuccessful school learners. The present study seeks to replicate the findings from these earlier studies using the longitudinal method.

Ss were first, second, and third grade children (N = 36) from middle class backgrounds and a like number from the same grades from lower class homes. A four level auditory memory test, consisting of CVC trigrams, nouns, anomalous sentences, and meaningful sentences, was administered individually to the children in 1967 and again in 1968.

A three factor ANOVA was used to analyze the data, with socioeconomic status, grade level, and auditory memory task level as factors. Significant differences in the performance of the two socioeconomic groups in favor of the middle class children were obtained. An interaction between task levels and socioeconomic status was observed. There were greater differences between the socioeconomic groups when language structure could be used to aid recall.

Analysis of the gain scores reveals an interaction between socioeconomic status and grade level. Gains are greater for younger, lower socioeconomic status children than they are for the older members of this group. Trends in the present study show the two socioeconomic groups becoming more similar with age up to seven years. The groups become dissimilar thereafter, with the older lower socioeconomic status children's performance gaining minimally during this one year lapse.

LANGUAGE DEVELOPMENT "GAMES" FOR CULTURALLY DEPRIVED CHILDREN, *Mrs. Cathy Kielsmeier* and *Jack Crawford*, Division of Teaching Research, Oregon State System of Higher Education

Acquiring effective communication skills is particularly difficult for the disadvantaged child from a bilingual or foreign language background. Typical language development programs developed for the monolingual disadvantaged child tend to be highly teacher centered and controlled, dependent on a trained specialist, and poor in providing prompt individual reinforcement.

The present study investigated the effects of language development techniques using a learning game approach designed to maintain (1) the initiation and sustenance of a high motivational level, (2) the frequent emission of linguistic responses, (3) immediate and dramatic feedback as to correctness of response, (4) the utilization of language as a tool to solve problems, (5) the development of language-based conceptual frameworks, and (6) effective instruction employing only teacher-aides or older children as instructors.

First-grade Mexican-American children were randomly assigned to one of three groups. The Experimental Language-game Group was exposed to the language development games for a 20-30 minute period daily for two school months. Control Group A was exposed to a series of games whose purposes were not language development. Control Group B received their usual classroom instruction only. All groups were given pre- and post-tests consisting of the Peabody and of an expressive language test developed for this study. This latter test consisted of four parts: (1) identifying and describing pictures of common objects and events, (2) giving comprehensible requests to another person, (3) placing a series of objects or pictures in ordered array and describing the array and its dimensions, and (4) giving comprehensible verbal indication of missing elements in this array.

Results indicated superior performance of the experimental language-game group across all measures.

Implications of the method for instruction in language and for research in the conceptual repertoire and strategies of the disadvantaged learner are discussed.

RELATIONSHIPS BETWEEN SOCIAL CLASS AND PHONEMIC AND NONPHONEMIC AUDITORY DISCRIMINATION ABILITY, *Thomas Oakland*, The University of Texas

This study examined the relationships between social class membership and performance on phonemic and nonphonemic auditory discrimination tests.

Three socioeconomic groups (upper-middle class, upper-lower class, and lower-lower class) of twenty Ss each were administered a phonemic auditory discrimination test and nonphonemic auditory discrimination tests of intensity, frequency, and pattern. On the phonemic auditory discrimination test, the upper-middle class (UM) group and the upper-lower class (UL) group performed significantly better than the lower-lower class (LL) group. On the nonphonemic auditory discrimination tests, the UM group performed significantly better than did the UL or LL groups on seven of the twelve measures. On no measure was a lower socioeconomic group significantly better than the higher socioeconomic group.

These results indicate that programs designed to remedy inadequate auditory discrimination ability should take into consideration possible nonphonemic auditory discrimination abilities.

THE EFFECTS OF STANDARD DIALECT TRAINING ON NEGRO FIRST-GRADERS LEARNING TO READ, *Richard Rystrom,* University of Georgia

This research is the first in a series designed to probe the relationships between the dialect spoken by Negro first graders and the problems they experience in learning to read. The following hypotheses were tested: (1) in eight weeks, Negro children can be taught to use elements of standard English dialect which do not occur in their native dialect; (2) the knowledge of this additional dimension of dialect will have a positive and significant influence on word reading scores; (3) dialect training will have a positive and significant influence on word reading tests in which the relationship between letters and sounds is controlled.

Two classes were chosen, and it was agreed that each teacher would continue to teach reading as she normally did. The classes were divided randomly into two groups. Half of one teacher's class was joined with half of the other teacher's class to form the experimental group. The students in the remaining halves were joined to form the control group. The treatment time chosen did not intrude on the time set aside for reading. The treatment group was given dialect lessons, and the control group was read stories. The following test instruments were used: (1) Rystrom's *Dialect Deviation Test,* (2) Gates' *Word Pronunciation Test,* (3) Fry's *Phonetically Regular Words Oral Reading Test,* and (4) the *Stanford Achievement Test.*

At the end of the treatment period, a two-way analysis of variance was conducted in order to measure teacher effect, treatment effect, and the interaction between them. The treatment F-scores were not statistically significant. All three hypotheses were rejected.

Session 17.8 **Division D**

RESEARCH ON TEACHER BEHAVIOR

Victor W. Doherty, Portland Public Schools, Chairman

RELIABILITY OF OBSERVATIONS OF TEACHERS' CLASSROOM BEHAVIOR, *Theodore Abramson,* City College of New York

Innovations in teacher training are often dependent on observational data. The problem of measuring reliability of observations collected by a team is due to (1) the difficulties of maintaining an observer team intact over an extended period of time and (2) observing each teacher more than once. These two conditions are normally required if one is to apply the Analysis of Variance (ANOVA) model proposed by Medley and Mitzel.

This paper takes these two conditions into account and presents a model which permits the calculation of an overall reliability coefficient and the partitioning of the sources of variation. That is, the model is designed for the observer team situation in which the team visits a number of different teachers only once and where the team does not necessarily contain the same members for all visits. The paradigm is developed for situations in which there are *n* observations per item per observer and also for the situations when there is only one observation per item per observer.

To apply the model adequately, techniques of identifying "Random" and "Fixed" variables are outlined. The model is applied to data derived from the use of an observational schedule by three teams. The schedule contains seven categories of teacher and pupil classroom behaviors. The overall reliability coefficient of .37 and the variance components of .38, .18, and .07 for the items, interaction, and error terms, respectively, indicate that the teacher and item factors account for 75% of the total variance. The model may also be used as a measure of the homogeneity of the items constituting an observational schedule.

OBSERVER EFFECTS ON TEACHER BEHAVIOR, *Thomas Samph*, Syracuse University

This study was undertaken to determine whether the presence of an observer has an effect on the verbal behavior of teachers and, if so, what the nature of that effect is. The following effects were predicted. (1) Teachers under observation would change their behavior to become more like their perceived ideal teacher. (2) If given prior notification that an observation was to occur, teachers would behave more like their perceived ideal teacher. This effect is predicted to occur independent of an observer's presence. (3) When being observed, teachers low in manifest anxiety would behave more like their perceived ideal teacher than would teachers high in manifest anxiety. The dependent variables in this study were five behavioral measures extracted from an interaction analysis matrix.

Ten female teachers were subjects in this study. They were observed under each of the following four experimental conditions:

1. Teachers not informed of an observation and no observer present in the classroom (normal classroom condition);
2. Teachers informed of an observation prior to its occurrence and an observer present in the classroom;
3. Teachers informed of an observation and no observer present in the classroom;
4. Teachers not informed of an observation prior to its occurrence and an observer present in the classroom.

In order to implement the design of this study it was necessary to employ a classroom monitoring system. Flanders' System of Interaction Analysis was used to code the recorded teacher verbal behaviors. A scale entitled "My Ideal Teacher" and Taylor's Manifest Anxiety Scale were administered to the teachers. These scales measured, respectively, a teacher's perceptions of an ideal teacher and a teacher's level of manifest anxiety.

Analysis of the data demonstrated the existence of an observer effect. Teachers became more "indirect" when an observer was present in their classroom, whether they were informed of an observation prior to its occurrence or not. The teachers in this study used more "praise" and "acceptance of student ideas" and less "criticism" when being observed.

Prior notification of observations by itself had no effect on a teacher's classroom performance. When prior notification was followed by an observer's presence, teachers decreased their use of "criticism." The level of a teacher's manifest anxiety was unrelated to the experimental conditions in this study.

Analyses of additional interaction variables supported the major conclusion of this study. The presence of a classroom observer leads to changes in a teacher's verbal behaviors.

ESTABLISHING RELIABILITY AND VALIDITY ESTIMATES FOR SYSTEMATIC CLASSROOM OBSERVATION, *Jeaninne Nelson Webb*, University of Alabama; *Bob Burton Brown*, University of Florida

This paper reports data obtained in a study designed (1) to compare two types of reliability, using both trained and untrained observers, in the observation of teachers' behavior, and (2) to explore the relationship between observer reliability and the validity of their systematic classroom observations.

Thirty-two observers, sixteen trained and sixteen untrained, repeatedly observed filmed teaching episodes over a period of time using the *Teacher Practices Observation Record* (Brown, 1968) as an instrument to systematically measure teachers' classroom behavior. Observation scores of the filmed teaching behavior were used in analysis of the data.

This paper focuses primarily on the development and assessment of two types of reliability estimates, between-observer and within-observer, and on a criterion-validity estimate of the observations of teaching behavior. Using multiple linear regression analysis, data were analyzed to identify variables, including training, which contributed to the reliability and validity of systematic observations.

The importance of this study lies in its head-on approach to establishing effective methods for the assessment of data collected by systematic observation. Throughout the literature on observational systems, the primary emphasis has been on achieving uniform agreement between observers as the method of procuring reliable data. If the question of the validity of observations has been raised, it has not been reported. Our findings indicate that between-observer reliability coefficients alone do not provide the information needed in order that confidence may be placed in data collected by observation. However, within-observer reliability, or the consistency of an observer's responses over time, was found to have a positive relationship with the validity of observations. The establishment of within-observer reliability and observer validity would seem to be necessary elements in the development and use of observation systems.

A METHOD FOR IDENTIFYING THE EFFECTIVE TEACHER, *Thomas B. Justiz*, University of California at Los Angeles

Can a reliable measure of general teaching ability be developed from pupil achievement scores in two different subject fields?

Ten student-teachers in one senior high school and 7 in another were evaluated according to their pupils' achievement scores on problem solving objectives, first in News Story Structure Concepts, then in Punched-Card Computer Concepts. All student-teachers were unfamiliar with both subjects and were supplied with a packet containing the two subjects in 'Kit' form. Each 'Kit' contained an objective, related subject matter and practice exercises, with instructions for the student-teachers to prepare the two lessons overnight. On the following day, each student-teacher randomly selected 18 experimental pupils from his training-teacher's class (12 pupils in the second school), and escorted them to a testing area. Pupil groups were then reconstituted, and each student-teacher was assigned to the one classroom which did not contain any of his own pupils. All student-teachers instructed, without supervision, for 30 minutes in each subject, then were given paper and pencil post-tests and 15 minutes for testing, under close supervision. (Pupils were again reconstituted between lessons at one school.)

Each student-teacher was ranked according to the mean score of his class in each subject field. The two rankings were then correlated, using the Spearman Rank-Difference Correlation Method. The correlations were statistically significant at the .05 level of confidence at both schools (i.e., teachers who were effective in one subject were usually as effective in the second subject as well). This appears to be the first reliable measure of general teaching ability.

The student-teachers were also ranked according to their scores on the Minnesota Teachers Attitude Inventory (MTAI). The MTAI rankings were correlated with both subject rankings, and were significant at the .05 level of confidence at both schools. The MTA thus appears to be a reliable predictor of the student-achievement producing abilities of student-teachers.

Session 17.9 **Division C**

CREATIVITY NO. 2

Gary A. Davis, University of Wisconsin, Chairman

AN EXPERIMENT IN THE DEVELOPMENT OF CREATIVE THINKING AND ITS EFFECT ON SELECTED COGNITIVE, AFFECTIVE AND PSYCHOMOTOR BEHAVIORS OF STUDENTS IN HIGH SCHOOL INDUSTRIAL ARTS COURSES, *Walter S. Mietus*, University of Maryland

Technically inclined boys whose interests, aptitudes, and abilities are not concomitant with traditional school subjects and academic expectations have need for creative expression. This study was an effort to meet this need and it included the following four major phases:

1. The development of a course of studies designed to include the range of mental operations defined by J. P. Guilford as central to creative thinking, five techniques for eliciting creative ideation, and practical lessons on the psychology of thinking.
2. The selection and training of seven teachers in the content and methods of providing an environment guided by C. R. Roger's theory of psychological safety and freedom.
3. The refinement of a technical content test of creative thinking and establishment of its validity and reliability. Three factors—uniqueness, quality and flexibility—were included in the instrument. Creativity was defined as a behavior distinguished by the capacity to produce a quantity of socially useful and original, or at least statistically infrequent, solutions to problems related to reality at a given time.
4. The completion of a sixteen week experiment using control (CG) and experimental groups (EG) and a battery of pre and posttests to determine the effectiveness of the program and its effect on

selected cognitive, affective and psychomotor behaviors. The relationship of creativity to these variables was also studied.

Treating the data with the analysis of covariance statistic it was found that:

1. The EG made significantly ($p < .01$) higher scores on the test of creative abilities, whereas the CG remained at an arrested level.
2. The EG scored significantly ($p < .01$) higher scores on the California Psychological Inventory scales of Femininity, Flexibility, and Dominance. The CG scored significantly higher on scales of Self-acceptance and Conformance. No difference was found on the scale of Independence. The evidence tended to support the hypotheses that personality characteristics and affective behaviors can be changed when boys are allowed to express their creative potential.
3. Significant ($p < .05$) positive correlations were found between creative abilities and CPI scales of Femininity, Flexibility, and Independence, whereas negative correlations were found with scales of Dominance, Conformance, and Self-acceptance.
4. Psychomotor abilities of the high and the low creative groups correlated negatively, whereas the average creative ability group correlated positively ($p < .05$).

This experiment provided positive evidence indicating that courses in industrial arts designed to elicit creative expression have far more desirable effects on cognitive and affective behaviors of boys than courses designed on the basis of industrial processes and products.

IDEA CHECKLISTS IN CREATIVE PROBLEM SOLVING: OVERCOMING RESISTANCE, *William E. Roweton* and *Gary A. Davis*, University of Wisconsin

The present study was one of a series of laboratory experiments which investigated the effects of idea checklists upon the creative performance of college age subjects who were asked to think of changes or improvements for a product (e.g., a kitchen sink). Early experiments in this series unexpectedly showed that subjects who were provided with lengthy, detailed idea checklists containing different colors, materials, shapes, etc., did *not* produce significantly more ideas than control subjects who were not provided with a checklist. Sometimes, the control subjects produced slightly *more* ideas than the checklist subjects. Apparently, a detailed idea checklist which gives problem solutions to the *S*, if only he will transfer the ideas to his score sheet, simply did not initiate a motivated flow of creative, associative behavior.

In the present experiment, subjects in the Checklist Group received the following exceedingly brief idea checklist, which was entitled "Aids in Thinking of Physical Changes": (1) add and/or subtract something, (2) change color, (3) change materials, (4) change by rearranging the parts, (5) change shape, (6) change size, (7) change design or style. For both the Checklist and Control subjects, the instructions simply read, "List as many physical changes as you can for a thumbtack (or kitchen sink)."

The results indicated that, first, the Checklist subjects produced two and one-half times the number of ideas generated by subjects in the Control Group. Second, subjects in the Checklist Group produced about five times as many ideas rated above the midpoint of a 7-point "creativity" rating scale.

The implication of these results is that, with college-level thinkers, an effective idea checklist must stimulate or challenge idea production. The abbreviated 7-item idea checklist accomplished this goal by providing subjects with general categories of problem solutions, which stimulated subjects to produce their own, specific ideas.

A COMPARISON OF THREE PROCEDURES FOR STIMULATING CREATIVE IDEAS, *Thomas F. Warren* and *Gary A. Davis*, University of Wisconsin

The present study compared the effectiveness of three techniques for stimulating idea production. The problem, given to 30 undergraduate subjects (*S*s), required them to "think of ways to change or improve a door knob." While such a task is normally completed in a fixed 10- or 20-minute period, present *S*s were allowed unlimited time. Thus, subject-determined problem-solving time, plus several measures of creative output, were the dependent measures.

One group (Short Checklist) was taught to use a seven-item checklist which suggested general categories of ideas (i.e., change colors, shape, material, design/style, size; rearrange parts; add/subtract something). A second group (Long Checklist) was given an expanded checklist, containing specific names of colors, patterns, materials, shapes, etc. *S*s in the third group (Checkerboard) were instructed in the Checkerboard procedure, in which ideas for improving one quality of an object (e.g., color or shape) are listed along one axis, and ideas for improving other qualities are listed along other axes. New ideas are produced by combining ideas on various axes. The fourth group (Control) received no instruction regarding techniques.

Each *S* received a pamphlet containing (a) an explanation of the particular technique, (b) the problem statement, (c) response sheets.

The major results were: (1) the four groups did not differ significantly in the amount of subject-determined problem-solving time, although *S*s in the Short Checklist Group used more time than *S*s in the other groups (supporting previous findings that short checklists are highly stimulating and challenging); (2) *S*s in the Checkerboard and Short Checklist Groups produced twice as many ideas as *S*s in the Long Checklist or Control Groups ($p < .01$); (3) Checkerboard and Short Checklist *S*s produced significantly more high quality ideas than *S*s in the other groups ($p < .05$).

A PROGRAM FOR TRAINING CREATIVE THINKING: THEORY AND PRELIMINARY RESULTS, *Gary A. Davis* and *Susan E. Houtman*, University of Wisconsin

This report describes (1) a three-part model of creativity, (2) a 6th-8th grade training program based upon this model, and (3) the results of the program's first field test. A "creative" student must possess, first, appropriate *attitudes* conducive to creative productivity, particularly a favorable attitude toward highly imaginative problem solutions, the attitude that his own creative problem solving skill can be increased, and an attitude of "constructive discontent," i.e., the notion that virtually anything can be changed for the better. Second, the student must possess cognitive *abilities* which facilitate whatever mental abstracting, combining, filling in gaps, perceiving, or associating contribute to the fluent production of original ideas. Third, the student needs a repertoire of *techniques* for the deliberate, systematic production of new combinations of ideas, such as looking for analogies and metaphors, identifying and improving individual attributes of a problem object, and other means of "forcing" original combinations of ideas.

The program, *Thinking Creatively: A Guide to Training Imagination,* incorporates concepts and principles from this three-part model by attempting to increase students' awareness of and appreciation for novel ideas, teaching techniques for producing new idea combinations, providing exercises for some creative abilities, and, through humor, creating a free atmosphere encouraging spontaneity and imagination. Generally, the program is in the form of dialogue among four characters, one of whom, a backyard scientist-inventor, teaches the other three creative attitudes and various problem-solving techniques.

In three idea-generating tasks, e.g., inventing new kinds of hot dogs, 21 7th grade students, who used the program as a text in a 10-week creative thinking course, produced an average of 50 percent more ideas than 32 comparable control students. The ideas of the trained students were (blindly) judged both "more original" and "better" than those of the control students. An accompanying attitude questionnaire indicated the trained students had developed favorable attitudes toward highly imaginative thinking, confidence in their own creative thinking ability, and an increased awareness of creative innovation.

Session 18.5 **Division C**

INTERACTION ANALYSIS NO. 2

David Starks, University of Michigan, Chairman

AN ANALYSIS OF TEACHER QUESTIONS, *Peggy J. Amidon,* San Francisco, California; *Edmund J. Amidon,* San Francisco State College

The purpose of this study was to investigate the types of questions asked by direct and indirect teachers and to determine the relationship among different types of questions and the cognitive levels of the students' responses that followed. The amounts and duration of pupil verbal participation following each of the question categories were studied. Thirty teachers in an elementary urban school were observed using Interaction Analysis. Audiotaped recordings of the most direct and the most indirect teachers for successive grade levels were then categorized according to an expanded Interaction Analysis system that incorporates Gallagher's and Aschner's four levels of questions and four corresponding levels of student responses. Student-initiated talk comprised the ninth category analyzed in this study.

Although the two groups differed in the total number of questions asked (direct 159, indirect 250), the number of student responses were similar (direct 378, indirect 385). The imbalance between the number of questions and responses in the direct group resulted from the large number of student responses following directions rather than questions. Both the percentage of student-initiated talk (direct 1.6%, indirect 36.3%) and the length of individual responses was greater for the indirect group. Student responses following divergent and evaluative questions were longer (divergent 4.92 seconds, evaluative 18.25 seconds) than that following factual and convergent questions (factual 3.54 seconds, convergent 3.45 seconds). The percentage of occurrences for each of the four types of questions differed significantly between the two groups (direct: factual 98%, convergent 2%; indirect: factual 42%, convergent 34%, divergent 5%, evaluative 18%).

This study has implications for future classroom research, future training programs concerned with teaching skills and those new curricular materials requiring varied questioning behaviors to achieve responses of varying cognitive complexity.

AN ANALYSIS OF PRESCHOOL TEACHERS' SPEECH STYLES USING THE COGNITIVE STIMULATION CODING CATEGORIES (CSCC), *Carla F. Berry, Virginia C. Shipman,* and *Robert D. Hess,* University of Chicago, Head Start Evaluation and Research Center

This paper reports the results of a pilot project to describe teaching styles by means of a new observation-coding scheme, the Cognitive Stimulation Coding Categories. The aim of the pilot study was, first, to document the presence of consistent, identifiable speech characteristics (styles) based on our selection of relevant measures. These measures cover Cognitive, Control, and Affective speech in interaction sequences. The second aim was to determine whether these speech characteristics provided a good measure of teacher differences.

Seventeen transcripts of teacher's speech were recorded and coded by trained coders. Reliability studies were completed and found adequate. An analysis of Variance and a Discriminant Function were performed. The statistical data from the pilot sample indicate the feasibility of the technique both for describing teaching styles and differentiating among teachers. The application of the coding scheme for future research is discussed.

This research touches three areas of current interest:

1. Development of techniques for describing the ongoing teaching process, especially with the focus of teacher-pupil interaction;
2. Cognitive development of preschool children and the relevance of language to cognitive growth;
3. The unique problems associated with research in a preschool setting.

THE EFFECTS OF INTERACTION ANALYSIS FEEDBACK ON THE VERBAL BEHAVIOR OF STUDENT TEACHERS, *Joseph Bondi, Jr.,* University of South Florida; *Richard Ober,* University of West Virginia

The purpose of the study was to investigate the effects of interaction analysis feedback on the verbal behavior of student teachers. Forty randomly selected elementary education students were trained in interaction analysis. In the quarter following the training, the forty students were observed weekly while they were engaged in student teaching. Data were collected from systematic observations conducted by four trained observers using a thirteen category modification of the Flanders System of interaction analysis.

Twenty of the subjects (feedback group) received weekly matrices and information sheets, while twenty (nonfeedback group) did not receive such feedback. After eight weeks of observations, data collected were analyzed by means of a Lindquist Type I Analysis of Variance. F-tests for group means difference were computed for twenty-four selected dependent variables.

The independent variables of the study were (a) the feedback provided the experimental group and (b) the lack of feedback in the control group. The dependent variables were selected observed verbal behaviors of student teachers.

Results of the study show that student teachers receiving interaction analysis feedback differed significantly from student teachers not receiving feedback in their use of the following teacher verbal behaviors: (1) they used more praise; (2) they accepted and clarified student ideas more; (3) they used more indirect teacher talk as opposed to direct teacher talk; (4) they used more extended praise; (5) they had more extended use of student ideas; (6) they used more positive affective talk; (7) they accepted student ideas more after teacher-initiated student talk; (8) they used more positive reinforcement after teacher-initiated student talk; (9) they used less corrective feedback; (10) they criticized students less; (11) they asked more questions; (12) they used less lecture; (13) they gave fewer directions. There was also less teacher-initiated talk and more student-initiated student talk in the feedback group.

COMPUTERIZED ANALYSIS OF BEHAVIORAL PATTERNS IN A FLANDERS INTERACTION ANALYSIS MATRIX, *Walter Doyle* and *Manfred Prokop*, University of Notre Dame

The purpose of this research was to generate a set of prescriptive statements for the sequencing of pattern constituents within the Flanders Interaction Analysis Matrix. This set was then utilized as an explicit guide for establishing and reading behavioral patterns in the matrix.

This set of explicit statements has six advantages: (1) consistency of analysis in the same and in different matrices; (2) avoidance of a reader's bias in establishing the patterns; (3) exhaustiveness in determining the pattern(s) within a single matrix; (4) accuracy of information as to the number of occurrences of a certain pattern; (5) standardization of selection of certain cells as "loaded cells"; and (6) greater comparability of two or more matrices of the same teaching situation. Although the standardizing guidelines were set up for computer use, they can also be applied in analyzing a matrix without the help of the machine.

A Pattern Variability Index (PVI), i.e., the total number of tallies in the transition cells of a given pattern divided by the number of all tallies utilized in that pattern, has been devised to determine the actual amount of interaction between teacher and students regardless of the time spent on prolonged use of "steady-state" behaviors. The smaller the Index the less change in behavior, and the less significant the pattern was if the criterion in analysis is the amount of actual interaction. This criterion has special significance for training prospective teachers in the planned use of behavioral skills.

Several implications for the analysis of matrices are discussed, such as the rationale for choosing the 5-5 cell as the beginning and the end of the pattern, and the relative significance of the steady-state cells.

Session 18.6 **Division C**

SPECIAL EDUCATION

Melvyn I. Semmel, University of Michigan, Chairman

INPUT-OUTPUT MODES AND SHORT-TERM MEMORY FOR OBJECT SEQUENCES IN GRADE I CHILDREN, *Peggy R. Koopman,* The University of British Columbia

Many studies which focus upon diagnosis and remediation of learning disabilities have attended to the mode in which material is presented or becomes available (input), and to the form in which the response is made (output).

The present study tests the proposition that, important as input and output modes may be separately in accounting for children's performances, there are children for whom input and output modes operate interactively. The input modes used here are Auditory, Visual, and Haptic (tactile-kinesthetic), and the output modes are Vocal and Motor.

The task chosen for the experiment was one of memory for object sequences. It involved presenting, under each of three input conditions separately, sequences of familiar objects increasing in number from two to eight. Following each sequence, every subject was required to indicate the objects that had been presented. This was done independently under both vocal and motor output conditions. The subjects were 90 Grade I children. Three sequences were presented at each level and there were two replications.

Derived scores were viewed as entries in a four-factor experiment having two fixed factors and two random factors. Standard ANOVA procedures revealed highly significant Subjects x Input x Output, and Subjects x Input interactions, and a significant Subjects x Output interaction. Estimates of variance components associated with each of these effects show the S x I x O interaction to account for about as much variance as the remaining interactions. The potential importance of Input x Output combinations in learning diagnoses seems clearly indicated.

Certain children, when viewed individually, were found to have performed particularly well or badly under specific combinations of input and output modes that seemed unrelated to whatever input or output strengths or weaknesses they had. Implications of this interactive role are discussed.

SOME EFFECTS OF VARIED EDUCATIONAL PLACEMENT FOR EMOTIONALLY DISTURBED CHILDREN, *Robert V. Turner,* Richmond Public Schools; *James D. Beaber,* Louisiana State University

The purpose of this study was to compare some effects of placing primary emotionally disturbed children in special education classes separate from secondary emotionally disturbed, or brain-injured, children.

The study employed a pre-test-post-test control group design. The experimental group consisted of two classes of children (N = 12) who had been diagnosed as emotionally disturbed and programmed together for educational purposes, and two classes of children (N = 14) who had been diagnosed as brain-injured and programmed together for educational purposes. The control group consisted of four classes of children (N = 25) who had been diagnosed as "learning disability," each class containing approximately equal proportions of emotionally disturbed and brain-injured children. These were programmed together or combined for educational purposes. The original sample contained fifty-one children with a chronological age range of 6 through 13. All groups received the educational program as provided in special education classes conducted in three public school systems. The experimental treatment consisted of separating pupils for instruction by etiology on the basis of clinical diagnosis.

Effects of separate placement were measured by the Slossen Intelligence Test, Gates-MacGinitie Reading Tests, Vineland Social Maturity Scale, A Class Play, Peer Rating and Self Rating, and an arithmetic sub-test of the Stanford Achievement Test.

Computerized variance and covariance techniques were employed to control variables found to affect relationships. Significant differences were found. Implications for intellectual and social functioning related to separate class placement were suggested.

A STUDY OF EMOTIONALLY DISTURBED CHILDREN IN REGULAR AND SPECIAL CLASSES, *Nicholas A. Vacc,* State University College at Fredonia, New York

This study was conducted to measure change in achievement and overt behavior of emotionally disturbed children in special and regular classes. In addition, the social position of emotionally disturbed and normal children in regular classes was assessed.

The study was conducted in public schools with children from centralized school districts. The population consisted of two matched groups of emotionally disturbed children attending regular and special classes. The group of emotionally disturbed children in the regular classes was compared with the children in the special classes in achievement and overt behavior. Additionally, the social position of the emotionally disturbed children in the regular classes was examined in relation to the normal children who were members of the same classes.

The Wide Range Achievement Test, Behavior Rating Scale, and a sociometric questionnaire were administered at the beginning and at the end of the study. A comparative description was made on the basis of the analyzed data and the following conclusions were made within the stated limitations of the study:

1. The emotionally disturbed children in the regular classes achieved less well on the Wide Range Achievement Test than the emotionally disturbed children in special classes.
2. While the emotionally disturbed children in the special classes made changes in overt behavior in a positive direction, the emotionally disturbed children in the regular classes showed change in overt behavior in a negative direction as measured by the Behavior Rating Scale.
3. The emotionally disturbed children in the regular classes were less well accepted than the normal children.
4. The emotionally disturbed children in the regular classes were more rejected than the normal children.
5. An analysis of the results of the sociometric questionnaire for stars, isolates, and rejectees indicated that (a) the percentage of stars was greatest among the normal children, (b) the percentage of rejectees was greatest among the emotionally disturbed children, and (c) the percentage of isolates was greatest among the emotionally disturbed children. Although the percentages varied between the fall and spring sociometric results, the data were consistent.

COMPREHENSION AND IMITATION OF SENTENCES BY MONGOLOID CHILDREN AS A FUNCTION OF TRANSFORMATIONAL COMPLEXITY, *Diane E. Greenough* and *Melvyn I. Semmel*, University of Michigan

The effect of sentence complexity on the ability of 40 Trainable Mentally Retarded (TMR) mongoloid *S*s ($\overline{X}_{IQ} = 34.3$) to comprehend and imitate verbally presented strings was examined in two studies. Four pairs of pictures were used as stimuli with each picture in a pair depicting one aspect of a reversible situation. Comprehension ability was tested by asking *S* to indicate which one of a pair of pictures was being described in the sentence spoken by *E*. Eight stimulus sentences were spoken by *E* for each pair of pictures: simple declarative or kernel (K), negative (N), passive (P), and negative passive (NP) for each picture. *S*'s ability to imitate sentences of four levels of complexity was tested by having him repeat the sentences used in the comprehension task, with passive forms shortened by deletion of the "by" clause so that stimuli would be of similar lengths.

*S*s correctly comprehended K sentences significantly more often than chance (mean percent correct = 58.7), but they correctly comprehended N sentences less often than would be expected by chance (mean percent correct = 37.6). Responses to P and NP strings appeared to be made at random. Imitation of K sentences (strings whose base structure have been modified only by obligatory transformations) was significantly better than imitation of sentences to which optional transformations (negative and/or passive) were applied. No significant differences were observed in accuracy of imitation of strings to which optional transformations had been applied.

Results were discussed in terms of competence and performance variables which might affect the ability of TMR *S*s to deal with verbal stimuli. Differential familiarity with the sentence types, with the more complex forms seldom used by adults with TMR children, and differences in length of sentence types were felt to be aspects of sentence complexity which, along with transformational complexity, affect the ability of TMR children to comprehend and imitate sentences.

Session 18.7 **Division C**

MEANINGFUL VERBAL LEARNING

A. J. H. Gaite, University of Wisconsin, Chairman

THE IDENTIFICATION OF WORD MEANING FROM SENTENCE CONTEXTS: AN EFFECT OF PRESENTATION ORDER, *Paul R. Ammon* and *Jack A. Graves*, University of California, Berkeley

An experiment was performed to explore the usefulness of Werner and Kaplan's Word-Context Test as a research tool in further studies of sentence comprehension and retention. Fourth and fifth grade children listened to lists of sentences, with each sentence in a list containing the same nonsense word. The task was to assign to the nonsense word a meaning which would fit all sentence contexts in the list. If a meaning met this criterion, it was considered a correct response.

Preliminary data provided an estimate of the probability that a particular sentence, presented in isolation, would elicit a correct response. The sentence most likely to elicit a correct response in a particular list was called the "best" sentence. Two presentation orders were used—from best to worst and *vice versa.* Furthermore, there were two response conditions. In the multiple response condition, *S* responded after each sentence in a list, whereas in the single response condition, *S* responded only after hearing the entire list. The two presentation orders and two response conditions comprised a two-by-two factorial design, with four groups of 15 *S*s, balanced for sex and grade level. The dependent variable was the number of lists for which *S* produced a correct response after hearing the entire list.

Analysis of variance indicated that significantly more correct responses occurred with the best-to-worst presentation order. Number of responses and the interaction of this factor with presentation order were nonsignificant effects. Further analysis suggested that, with the worst-to-best order, early sentences interfere with normal processing of subsequent sentences. This effect may have implications for a model of sentence comprehension.

SOME VERBAL ASPECTS OF SCIENTIFIC CONCEPTS, *Thomas E. Curran, Paul E. Johnson,* and *David L. Cox,* University of Minnesota

Scientific concepts are posed most precisely in terms of the formal language of mathematics, but are often communicated through ordinary discourse. To be studied psychologically, the elements of verbal communication must be described; one means of describing these elements is to identify the responses which they produce.

This study deals with three classes of such responses: verbal associations, judgments of similarity, and verbal definitions. Subjects in the study were college seniors and graduate students in physics. Similarity in patterns of verbal associations to concept words in physics were taken as indices of agreement in the meaning of the terms across subjects. Similarity judgments and verbal definitions of these same words were then examined to study the generality of agreement in meaning across tasks. These results were compared with the outcome of previous studies where the subjects were students of high school physics.

Two major conclusions were reached. First, an apparently well-integrated verbal-theoretical structure is exhibited by the associations of the sophisticated students. The uniformity in patterns of association across *S*s in the earlier studies with less-sophisticated *S*s was not unexpected, since they must rely almost exclusively upon ordinary discourse for the meaning of terms. But one might predict that with increasing experience, verbal structures become more diffuse due to the extensive use of mathematical terminology. The uniformity among the associations of the more sophisticated subjects leads one to presume either that verbal structure does not develop much beyond the naive level, or that such structure is inherent in the science itself and is developed and refined with increasing mastery in the subject matter. Present data suggest that a well-developed verbal structure exists, and moreover, that there are differences in this structure between the levels of mastery.

A second conclusion relates to the robustness of the verbal structures themselves. High correlations were found among the responses across tasks, dissimilar as these tasks may be; something which suggests considerable stability in the verbal meaning of concepts in physics.

The presence of well-developed verbal structures at several levels of subject matter mastery and the robust nature of these structures across tasks supports a linguistic approach to the analysis of scientific concepts.

A COMPARISON OF TELEVISION AND AUDIO PRESENTATIONS OF THE MODERN LANGUAGE ASSOCIATION FRENCH LISTENING EXAMINATION, *William M. Stallings* and *Richard E. Spencer*, University of Illinois

One objective of modern foreign language instruction is an ability to converse with native speakers. In a standardized examination listening is usually tested by aural stimuli broadcast from an audio-tape. However, in face-to-face communication, contextual and non-verbal cues are available. Also in teaching, the instructor provides certain non-verbal cues. Hence it would seem that the testing of listening comprehension might be improved by using a video-tape which could offer non-verbal cues such as hand gestures and lip movements in addition to aural stimuli.

For this study a video-tape was made from the script of the Form MB Modern Language Association French Listening Examination. Native French speakers were selected as actors. An audio-tape was recorded from the video-tape.

The subjects were second year French students. This experiment was conducted twice, at the end of the fall and spring semesters. Sections of second year French classes were randomly assigned to auditoriums having TV or audio receiving capabilities. Unfortunately, during the experiment TV reception in several auditoriums was not optimal.

From six different comparisons, only one significant difference was found between audio and video mean scores: the audio score was higher. In the preceding comparison, the TV data were from an auditorium which experienced poor TV reception. However, in comparisons involving auditoriums with good TV reception, the mean video scores were not significantly higher than the mean audio scores. An attitude scale toward TV testing was administered to fall semester students who were tested by TV. Those students appeared to have slightly unfavorable attitudes towards TV testing. Perhaps, the quality of TV reception is critical to the success of TV testing. Nevertheless, no support was given to the hypothesis that providing non-verbal cues will increase listening comprehension scores.

Session 18.10 **Division D**

RESEARCH ON TESTS

Richard C. Cox, University of Pittsburgh, Chairman

AN EVALUATION OF THE D-48 (DOMINOES) TEST AS A CULTURE-FAIR MEASURE OF INTELLIGENCE, *Conrad G. Katzenmeyer*, Kent State University; *Edmund D. Thomas*, Naval Personnel Research Activity; *Irina Dale*, San Diego State College

The D-48, or Dominoes, is a non-verbal test that would appear to have considerable promise as a culture-fair measure of intelligence. While the test loads heavily on the intelligence or g factor, its content is equally familiar to members of a wide variety of sub-cultures. A number of European studies support the value of the D-48 as a culture-fair measure, but no investigations have been carried out with American sub-cultures. Indeed, there is very limited information of any type on the use of this instrument in the United States.

In this study, the D-48 was included in the classification battery administered to approximately 3000 Navy recruits in their first week of military training. Because of recent Defense Department programs designed to provide training for disadvantaged youths, a sizable sub-sample of these men were from culturally deprived backgrounds. The D-48 and other verbal measures of intelligence from the classification battery were correlated with Recruit Final Achievement Scores, an essentially academic criterion. Analyses were carried out both within and across racial, educational, and socioeconomic level sub-classifications, statistically controlling for regression to the mean.

Initial results support the hypothesis that the D-48 test does have some advantages over traditional intelligence tests when testing the culturally deprived. Further analyses exploring the interactions of D-48 scores and the various status variables for predicting academic achievement will be presented. In addition, a comprehensive set of test statistics for the D-48 test will be provided, with implications for its future use.

RELATIONSHIP BETWEEN REMOTE ASSOCIATES ABILITY AND THE DOGMATIC PERSONALITY, *David A. Johnson*, The Ohio State University

This study was an attempt to test the construct validity of Mednick's Remote Associates Test (RAT) (Mednick, 1962) as a measure of creativity outside the cognitive domain. It was noted that the open and closed mind construct (Rokeach, 1960) incorporates many of those personality characteristics correlated with high and low creativity criterion groups (Barron, 1955; McKinnon, 1962). Mednick constructed the RAT specifically to measure the ability to form new associations which he defined as creativity. If creativity as determined by these personality traits and creativity as measured by this cognitive measure have any common characteristics, one should expect a correlation between Hi-RAT scores and open-mindedness (i.e., a negative correlation between RAT and dogmatism scores.)

Dogmatism, RAT, and CMMI intelligence data were collected from 223 high school juniors and seniors. A comparison was made between the dogmatism means of the upper and lower thirds of the RAT distribution. A correlation coefficient was also computed between RAT and dogmatism for the entire sample. Intelligence and sex were controlled by matching in the comparison of means and by partial correlation in the correlation paradigms.

The hypothesis was confirmed at the .01 level in both paradigms. With intelligence controlled, a —.32 correlation ($p < .01$) was obtained from the entire sample. The by sex analysis of both paradigms demonstrated that the observed results were due to the male members of the sample.

This study supports the further use of the RAT as a measure of the cognitive aspects of creativity. In a more general sense, this study demonstrates that performance is the result of a complex interaction of cognitive, personality and other factors.

COGNITIVE AND PERSONALITY REFERENTS OF CREATIVITY: SOME RELATIONSHIPS, *Fred W. Ohnmacht*, University of Missouri

Three studies are discussed which explored the relationships among cognitive and personality referents of creativity. Implications of these studies for the identification of potentially creative Ss are presented.

Study I. Relationships between the 16 Personality Factor Questionnaire (16 PF), Barron Complexity Scale (BCS), and measures of divergent thinking from Guilford's Structure of the Intellect (SI) were examined utilizing correlational techniques. Previous work with the 16 PF, indicating discrimination among creative and non-creative groups, suggested that scales should be significantly related to the BCS and SI measures. Results indicated that the 16 PF and BCS demonstrate limited congruency, whereas there was essentially no correspondence across the cognitive and personality domain.

Study II was similar to Study I except that a broader sample of divergent thinking tasks from the SI were employed. Further, the Myer-Briggs Type Indicator (MBTI) was employed to supply personality referents of creativity. As with Study I, congruency between cognitive and personality referents was not observed in the data.

Study III examined the congruency of different personality referents of creativity. Specifically, 16 PF and MBTI referents were examined for congruence. As with the 16 PF and BCS results in Study I, a limited congruence as manifested by low but significant correlations was observed.

Although different samples were involved, the results can be viewed as an approximation of the Campell-Fiske multi-method, multi-trait approach to test validation. Cast in these terms, mono-method convergence (16 PF, BCS, MBTI) was moderate at best, whereas multi-method convergence was nil. Inter-domain correspondence (personality with cognitive referents) was such that extreme care should be taken in comparing studies which identify potentially creative Ss by employing only personality or cognitive measures. If the results reported here have generality, quite different sets of Ss will be identified as potentially creative if personality or cognitive referents are employed.

ANXIETY AND LEARNING TO FORMULATE HYPOTHESES, *Stephen P. Klein*, Center for the Study of Evaluation of Instructional Programs; *Norman Frederiksen*, Educational Testing Service

This study investigated some effects of a training procedure designed to increase the production of original ideas and the relationship of such behavior to certain personality and ability measures.

A sample of 127 college students was split into control and experimental groups. These Ss took a test battery that included measures of verbal fluency, verbal ability, "originality," test anxiety, and defensiveness. The only difference between the two groups was that the experimental Ss received training in the form of feedback information about their performance on one of the originality measures.

A series of analyses of variance and covariance indicated the following: (1) training increased the quantity but not the quality of the ideas produced, (2) this increase did not transfer to a similar task involving divergent production of ideas, and (3) test anxiety has a consistent curvilinear relationship with performance, poor performance being associated with an intermediate level of test anxiety.

The preceding results have a number of implications for both testing and teaching. From the standpoint of testing, it appeared that the treatment influenced the S's response set more than his ability. Similar measures of originality also might be susceptible to this bias. An analogous point is that attempts in the classroom to modify experimentally an S's creativity may only affect his attitude (in the direction of lowering his standards)

and not his ability or real achievement. Finally, the curvilinear relationship between test anxiety and performance on several measures (in a direction contrary to classical drive theory) has several implications for prediction and the use of moderator variables.

Session 18.11 **Division D**

STUDIES RELATING TO VERBAL PERFORMANCE

Charles A. Woodbury, The Macmillan Company, Chairman

A STUDY OF THE RELATIONSHIPS BETWEEN PRIMARY GRADE PUPILS LABELED AS EITHER CULTURALLY DISADVANTAGED OR CULTURALLY ADVANTAGED AND THEIR DEVELOPMENT OF CERTAIN LANGUAGE SKILLS, *Marian Lee Vick*, North Carolina Agricultural and Technical State University; *Milton D. Jacobson*, University of Virginia; *Joseph Carlton Johnson II*, Duke University

The purpose of this investigation was to study the interrelationships among the language skill complex, grade level, intelligence level, race, sex, father's occupation, school, chronological age, and general reading ability of children placed in the primary grades with respect to their cultural index.

The language skill complex consisted of auditory discrimination, articulation of speech sounds, recognition vocabulary, and vocabulary of use.

Cultural index was based on the subjective identification of children by their teachers as either culturally disadvantaged or advantaged.

Procedures consisted of the selection, administration, and interpretation of tests that purport to measure language skills at the primary grade level. The tests utilized were the *Institute for Personality and Ability Testing Culture Fair (or Free) Intelligence Test* (Scale 1), the *Culture Fair Intelligence Test* (Scale 2, Forms A and B), the *Gates-MacGinitie Reading Tests, Wepman's Auditory Discrimination Test* (Form 1), the *Templin-Darley Test of Articulation,* and a modified form of the *Seashore-Eckerson English Recognition Vocabulary Test.* Four open-ended questions were utilized to secure a sample of verbal responses. In addition, children were identified as being either culturally disadvantaged or advantaged by their teachers and were subsequently labeled as such for purposes of this study.

The test data were collected and subjected to analysis. Computerized analysis was accomplished through the variance and covariance techniques of multiple linear regression.

***S*s were 322 primary-grade children from two elementary schools. One school was classified as having enrolled a predominantly disadvantaged population, while the other was categorized as having enrolled a predominantly advantaged population.**

The primary null hypothesis examined in this investigation was that the development and utilization of the total language skill complex will not differ significantly between groups of primary grade-level children who have been subjectively labeled by their teachers as either culturally disadvantaged or advantaged. One subordinate hypothesis was tested: Any relationship between the development and utilization of the language skill complex by primary grade-level children is not dependent on any of the following variables: grade level, intelligence level, race, sex, father's occupation, school, chronological age, general reading ability, and cultural index.

Results indicated that the null hypothesis could not be rejected. In

testing the subordinate hypothesis it was found that grade division, age, intelligence level, and general reading ability were singly or in combination significantly related to certain specified factors constituting the language skill complex when these latter variables functioned as criterion measures.

AN EXAMINATION OF THE INTERDEPENDENCE AMONG VARIABLES RELATED TO SPECIFIC PREDISPOSITIONS AS MANIFESTED THROUGH POPULAR READING MOTIFS IN THE GRAMMAR GRADES, *Joseph C. Johnson II*, Duke University; *Milton D. Jacobson*, University of Virginia

The purpose of this investigation was to examine the reader in the upper grammar grades with respect to the manner in which his psychological set, as determined by Likert type attitudinal appraisal devices, toward specified motif content is related to his comprehension of reading selections reflecting such content. Specifically, the procedures for this study were such as to compare the literal and interpretative comprehension abilities of intermediate grade children on certain thematically based reading selections with their attitudes toward that material.

The investigation was designed to control the predictor variables of race, sex, intelligence level, school, grade, attitude, literal reading comprehension, interpretative reading comprehension, socioeconomic status, and chronological age and the criterion variables of general reading achievement level, anthropomorphic attitudinal set, underdog attitudinal set, and culturally-alien attitudinal predisposition.

Three hundred twenty intermediate grade children were initially tested with standardized and informal instruments to obtain the information necessary to control the variables specified above. Of these, one hundred fifty were retained for the study since the preliminary test data indicated that their intelligence levels, reading achievement levels, and attitudinal responses were distributed over the entire range of those traits.

Results indicate that the consociations between attitude and reading comprehension were not affected by race. Consequently, materials written primarily to develop or enhance reading comprehension for certain racial groups appear to offer little promise. The study also found that literal and interpretative reading comprehension are differentially affected by intelligence level, sex, and socioeconomic status. This finding raises an argument for providing different stories for various individuals representing sundry classifications of those categorical variable scales.

Additional results and implications are discussed in the paper.

VERBAL INTERACTION PATTERNS AS A FUNCTION OF DIFFERENT LEADERSHIP STYLES, *Anthony J. Reilly*, University of Michigan

The purpose of the study was to investigate the verbal behavior of leaders who were operationally defined and identified through the use of Fiedler's Least Preferred Coworker measure. Subjects were 86 nursing students from a large midwestern hospital. Fourteen small groups participated in ten different problem-solving sessions. Group observers independently classified the verbal interaction of the leaders into 18 behavioral categories. The experiment was conducted over a period of ten weeks.

There was no evidence that differently styled leaders varied in the number of responses made to specific individuals in their respective groups. Furthermore, there were no significant differences among types of leaders with respect to the extent to which they responded qualitatively differently (using different verbal categories) to different members of their groups. Also, there were no significant differences among the types of leaders in the variance concerning change in the number of responses made to various individuals in their groups.

It was found that task-oriented leaders (leaders who obtained low LPC scores) addressed their group members relatively early in the problem-solving sessions, whereas high LPC, relations-oriented leaders addressed their group members relatively later in the session. On the average, however, there was no evidence that the two types of leaders addressed their groups with different frequency.

It was found that, in general, task-oriented leaders interacted more with their group members then did relations-oriented leaders. Moreover, it was found that category 6 (gives procedural suggestions), category 14 (disagrees, maintains a contrary position), and category 19 (asks for repeat, clarification) were used more by the relations-oriented leaders. Judging from the content of these categories these results suggest that high LPC leaders were more directive and low LPC leaders were more supportive.

For this study, the behaviors of the leaders, as manifested through their verbal communication patterns, did not support Fiedler's differential descriptions of relationship-oriented and task-oriented leaders.

It is suggested that follow-up research include not only verbal behavior of leaders, but other overt behaviors in an attempt to better understand the meaning of "psychological distance." It is hypothesized that the superiority of the relationship-oriented leaders may be attributed to a higher-order organizational variable which outweighs the particular type of task being worked on. It is suggested that this factor of organizational climate be included in future research on the theory.

IQ STABILITY ON VERBAL AND NON-VERBAL INTELLIGENCE TESTS, *Kenneth D. Hopkins* and *Glenn H. Bracht,* University of Colorado

Bloom's (1964) compilation of research studies of IQ constancy with children revealed only four studies in which IQ stability was investigated for group tests. None of these studies reported the IQ constancy of non-language tests. Of the four studies in which group tests were employed, the greatest time interval reported was only five years. The present investigation explored the degree of stability and change of language and non-language IQ scores for a large sample of normal pupils from grade one through grade eleven.

First-grade pupils in a moderately large school district were followed through grade eleven, being given verbal and non-verbal group intelligence tests (CTMM and Lorge-Thorndike) in grades one, two, four, seven, nine, and eleven. Sample sizes for pairs of grades varied from 240 to 900. The degree of IQ stability between each pair of testings was investigated for both types of tests as well as for the total IQ scores.

The IQ stability on the non-language tests was consistently and often substantially less than for the language scales on both the CTMM and the

Lorge-Thorndike. IQ scales from group tests appear to have considerably less stability than do individual tests, especially for the primary grades where even one-year stability was quite low. Verbal IQ variance stabilized considerably at grade four, but the non-verbal scores reflected little lasting variance until grade seven.

Session 19.1 **Division C**

COGNITIVE DEVELOPMENT: PIAGETIAN CONSERVATION NO. 2

AN ANALYSIS OF THE EFFECTS OF SELECTED EXPERIENCES ON THE ABILITY OF PRE-SCHOOL CHILDREN TO USE THE TRANSITIVE PROPERTY, *Russell L. Carey* and *Leslie P. Steffe,* University of Georgia

The purpose of this study is to analyze the effects of selected experiences in length comparison, conservation of length, and conservation of length relations on the ability of four and five year old children to use the transitive property of length relations.

The subjects were 19 four-year-old children and 32 five-year-old children. All children received instruction in establishing length relations between two curves, conserving length relations, and conserving length. Small group instructional procedures were utilized. A transitivity test was administered after the Length Comparison Instructional Unit and after the Conservation Units. The pupils were treated on a one-to-one basis for evaluation.

The conclusions are summarized as follows:

1. Experiences in establishing length relations do not appear to be sufficient for four- and five-year-old children to use qualitative transitivity of length relations.
2. Formal experiences in establishing length relations, conserving length relations, and conserving length does increase the ability of five-year-old children to use qualitative transitivity. These same experiences do not seem to increase the ability of four-year-old children to use the transitive property.
3. There appears to be little, if any, relationship between the variables Age, Verbal Maturity, I.Q., and Social Class and the ability of four- and five-year-old children to use the transitive property.

SEMANTIC FACTORS IN CONSERVATION OF WEIGHT, *Susan G. Nunmedal* and *Frank B. Murray,* University of Minnesota

It has been shown that young children often confuse connotative and denotative meaning, and that this confusion is related to errors they make in the discrimination of some of the same physical attributes of objects that have been investigated in conservation studies. This investigation explored the relation between this connotative-denotative confusion and nonconservation of weight.

Twenty-seven first- and 30 second-graders were tested on a series of eight connotative-denotative discrimination problems and seven conservation of weight problems, each of which represented physical dimensions that are highly correlated on the semantic differential potency factor. In the discrimination problems subjects were given two objects that actually differed in only one dimension, such as weight, and were asked if the objects also differed in such other dimensions as hardness, smoothness, size, strength, etc. In the conservation of weight problems, after one of two identical clay balls was transformed along the dimensions used in the discrimination problems, subjects were questioned about the equality of two balls' weight.

A significant relationship between the ability to conserve weight and the ability to make denotative discriminations was found and indicated that the general ability to make correct denotative discriminations was acquired prior to the ability to conserve weight. It was concluded that the discrimination of denotative from connotative meaning is a necessary, but not sufficient, condition for conservation of weight. Significant differences in the proportion of conservers and nonconservers in both grades on each of the weight conservation problems were found, implying that conservation of weight is a complex and transformation specific concept and not a simple, all-or-nothing concept as Piaget has often maintained. Additional transformations which are psychologically relevant in conservation problems might be generated and discovered from physical dimensions correlated on other semantic differential scales.

A STUDY OF INTERRELATIONSHIPS OF CONSERVATION OF LENGTH RELATIONS, CONSERVATION OF LENGTH, AND TRANSITIVITY OF LENGTH RELATIONS AT THE AGES OF FOUR AND FIVE YEARS, *Leslie P. Steffe* and *Russell L. Carey*, University of Georgia

The purpose of this study was to investigate interrelationships among the ability of four- and five-year-old children to conserve length relations, conserve length relations involving properties of the relations, conserve length and use the transitive property of length relations at two different points in time: (1) after formal experiences only in establishing a length relation between two curves and (2) after selected formal experiences in (1) above and in conserving length relations and length.

The subjects for the study were 20 four-year-old children and 34 five-year-old children in the Suder Elementary School, which is an experimental school for the Research and Development Center for Educational Stimulation, University of Georgia. The children were in three self-contained classrooms with fours and fives in each room. Small group instructional procedures were utilized. All evaluation was conducted on a one-to-one basis. Interrelationships of the variables were studied via 2 x 2 contingency tables.

The instructional and evaluational sequence was: (1) Length Comparison Instructional Unit, (2) Conservation of Length Relations Test, (3) Conservation of Length Test, (4) Transitivity of Length Relations Test, (5) Conservation of Length Instructional Unit, (6) Conservation of Length Relations Instructional Unit, and (7) re-administration of tests (2), (3), and (4).

The conclusions of the study are as follows: (a) an ability to conserve length relations precedes an ability to use properties and consequences of length relations involving conservation of length relations; (b) an ability to conserve length relations and use properties and consequences of length relations precedes an ability to use transitivity of length relations; and (c) an ability to conserve length is not related to an ability to conserve length relations, use the transitive property of length relations, or use properties and consequences involving length relations.

INDUCING CONSERVATION OF NUMBER, WEIGHT, VOLUME, AREA, AND MASS IN PRE-SCHOOL CHILDREN, *Beverly S. Young*, Stephen F. Austin State College

The purpose of the study was to determine whether, by using a multivariate approach, conservation of number, volume, weight, area, and mass could be induced and retained in three- and four-year-old children. Other relationships which were noted were the correlation of factors comprising conservation of number, the correlation of various areas of conservation, how rote counting, rational counting and attentiveness correlated with conservation, and whether instruction influenced conservation more at one IQ, CA, or MA level than another.

132 preschool subjects were used, ranging in CA from 3-4 to 5-5, IQ from 85-166, mean 122. An individual pre-test, post-test, and S-B IQ test were administered. The criterion test, based on Piaget's tasks, consisted of two 58 item equivalent forms, reliability .96. The experimental group received 8 ten-minute instructional periods.

The findings showed highly significant gains in conservation in all areas: number, volume, weight, area, and mass. The gain was consistent across all CA, MA, and IQ levels. Subjects tested three weeks following instruction showed a slight increase in total conservation over those who were post-tested immediately following instruction. Therefore, it was concluded that the concept had been retained and somewhat further developed.

Correlations between CA, MA, IQ and conservation were moderate and positive. The correlations between rote counting and conservation and rational counting and conservation were moderately high. Rote counting correlated nearly as high as rational counting at this age level. Attentiveness also had a strong positive correlation with composite conservation score.

Session 19.2 **Division A**

BUREAUCRACY AND INNOVATION

Irwin T. Johnson, University of Wyoming, Chairman

A MODEL FOR EDUCATIONAL IMPROVEMENT, *Burton W. Kreitlow*, University of Wisconsin

The purpose of this research was to describe the process of educational improvement as it takes place in operating school districts and translate this process into a physical structure that can be shown by two and three dimensional models.

To describe the process and develop a viable model, the results of observations in ten Wisconsin schools from 1949 through 1967 were integrated with an analysis of a variety of models for change developed by social scientists in both agriculture and education. This, in turn, was examined against the backdrop of tape recordings of committees for change functioning within three experimental and five control communities from 1966 through 1968.

The development of the model was based on several operating assumptions. Among these are the following:

1. All school districts have an external and an internal structure which provides boundaries of process.
2. School districts have social machinery that can be used to effect change.
3. The improvement process can be mapped.
4. The public school social system tends to be inefficient.

A new Model for Educational Improvement has been developed and is available for demonstration in both two and three dimensional forms. It incorporates selected characteristics of previous models, notably those of Rogers, Bruce, Guba and Clark, and Gideonese. The process of educational improvement within a school district can be observed with considerable objectivity against the backdrop of the Model.

The value of this kind of a model to the practitioner lies in the number of things it will explain about how his school system operates. Initial testing of the Model for Educational Improvement led to the conclusion that it has real value. The administrator, teacher, board member, and citizen is now able to compare the process for improving education in his district with a model that gives form to that process.

BUREAUCRATIC CHARACTER OF ADULT EDUCATION ORGANIZATIONS AND INNOVATIVENESS IN PROGRAM DEVELOPMENT, *Wayne B. Ringer*, Utah State University; *William S. Griffith*, The University of Chicago

The purpose of this study was to determine the relationship which might exist between certain dimensional bureaucratic characteristics of administration and the innovativeness of an organization in program development.

The perceptions of staff members in forty-five Cooperative Extension Service organizations relative to five selected bureaucratic characteristics of administration were obtained through the use of a mailed questionnaire. These characteristics were compared with organizational innovativeness in program development over a five year period as observed and reported by fifty-three raters composed of extension administrators, one group responsible for programs within the states and the other group representing a national perspective.

The bureaucratic characteristics, "hierarchy of authority," "rules and procedures," and "interpersonal relations," did not prove to be significantly related to innovation when tested as independent variables by correlation and multiple regression analyses. The bureaucratic characteristics, "division of labor" and "rewards of office," were significantly related to innovation, and when the characteristics, "rewards of office" and "rules and procedures," were tested as grouped or combined variables rather than independently, they proved to be quite significantly related to innovation in program. Another variable, the amount of "human and material resources" possessed by an organization, was found to be highly significant as a predictor of organizational innovativeness in program. Approximately one-half of the variance in innovation among the adult education organizations studied was found to be attributable to the independent variable, "division of labor," and to the grouped variables, "rules rewards" and "resources."

Among other things it was concluded that the absence or presence of constraints in the forms of specified rules and procedures may be looked upon as part of the reward system, and that the reward system and the division of labor may influence the innovative output of an organization in program development.

THE RELATIONSHIP OF STAFF TENURE AND ADMINISTRATIVE SUCCESSION TO STRUCTURAL INNOVATION, *Paul P. Preising*, San Jose Junior College District, San Jose, California

This investigation sought to determine whether schools in different states of structural innovation also differed significantly on each of the following variables: length of tenure of staff, source of recruitment of administrators, size of school, size of district, and expenditure per pupil.

Structural innovations were defined as those innovations requiring changes in the elements which order the operations and functions of the entire school. Structural innovations were further distinguished from non-structural innovations by referring to changes that cut across classroom and departmental boundaries and that are not restricted to the content or organization of one or a few subject fields. The specific structural innovations selected for study were (1) use of teacher aides, (2) team teaching, (3) variation in class size, and (4) variation in length of class period.

Of schools categorized as high, medium, and low on structural innovation, it was hypothesized that there would be a significant difference ($P = .05$) in:

1. Mean length of tenure of superintendents;
2. Mean length of tenure of faculties;
3. Mean length of tenure of principals;
4. The ratio of outside to inside recruitment of principals;
5. The ratio of outside to inside recruitment of superintendents;
6. Mean expenditure per pupil;
7. Mean size of district;
8. Mean size of school.

The hypotheses of the study were tested separately in a sample of 105 Oregon and 309 California public high schools. The major hypotheses were tested using a one-way analysis of variance design: chi square techniques were used to test the minor hypotheses.

The first four hypotheses were not supported ($p > .05$). Hypothesis number 5 above received support for California ($p = .033$) but not for Oregon ($p > .50$).

Hypothesis number 6 above was supported in California ($p < .05$) but not supported in Oregon ($p = .18$).

The last two hypotheses (numbers 7 and 8 above) received support in both states ($p < .05$).

SCHOOL FACULTY MEETINGS—AN INTERACTION ANALYSIS, *David B. Crispin*, Indiana State University; *R. Duane Peterson*, Wayne State University

The purpose of this field study was to gather data relative to the question, "Does the behavior of the principal significantly affect the behavior of the teachers during school faculty meetings?"

The following hypotheses were tested:

(1) the *more indirect* the principal's behavior, the *more supportive* will be the teachers' behaviors (behaviors of individuals);

(2) the *more indirect* the principal's behavior, the *more supportive* will be the group behavior (G factor);

(3) the *more indirect* the principal's behavior, the *greater* will be the *amount of participation* on the part of the teachers (the higher will be the teacher/principal ratio).

Terms were defined as follows:

(1) *Direct behavior* (Principal)—the principal *uses his authority,* states his or the administration's position, ideas, or requirements; gives orders, directions—lectures, reads prepared directives or statements of policy; justifies his position of authority.

(2) *Indirect behavior* (Principal)—the principal *willingly shares his authority with the teachers,* shares and accepts both feelings and ideas, *seems genuinely willing to be influenced by the teachers,* encourages them to express themselves, aims for consensus.

(3) *Non-supportive Behavior* (teachers)—the teacher seems to be participating *only because he has to;* he is defensive, uncooperative; he seems *bored,* unattentive, uninterested; he is *perfunctory,* seems to be behaving simply to get rid of the task.

(4) *Supportive Behavior* (teacher)—the teacher seems *willingly cooperative;* spontaneously, enthusiastically contributes feelings and ideas; he is attentively, seriously involved.

(5) *Independent Variable*—the behaviors of the principals.

(6) *Dependent Variable*—the behaviors of the teachers.

(7) *Controls*—all faculty meetings were held after school and lasted 40-50 minutes. All principals and teachers discussed the topic, "Utilization of Teacher Time." All the schools are elementary.

Observers trained in interaction analysis (reliability = .90) recorded behaviors every three seconds or faster during 76 faculty meetings in the Wabash Valley, Indiana, and metropolitan Detroit school districts.

The major finding of the study was that, even though the principal's behaviors varied significantly, the teacher's behaviors were very much the same.

Session 19.3 **Division C**

CONCEPT LEARNING NO. 3

Martin R. Wong, Michigan State University, Chairman

THE EFFECTS OF CHILDREN'S DIMENSIONAL PREFERENCES UPON CONCEPT LEARNING, *Claude E. Hill, M. C. Wittrock,* and *C. Norwood,* University of California, Los Angeles

The study was designed to investigate the effects of dimensional preferences upon the transfer from learning a unidimensional concept to the learning of a bidimensional concept.

One hundred and forty-eight fourth grade students were randomly assigned to one of six treatment groups or a control group in a standard proactive transfer design. Concept identification training was given utilizing a card sorting task. A deck of 32 stimulus figures was generated from all possible independent combinations of five bi-valued dimensions.

A pre-test, administered prior to original learning (OL), indicated that an order of preferences existed for the five stimulus dimensions. During OL Ss were trained on either one or two unidimensional problems which were either relevant or irrelevant to the bidimensional transfer problem. The learning rates for the unidimensional problems were found to covary positively with the dimensional preferences established by the pre-test.

Following OL all Ss learned the bidimensional concept. Two weeks later a retention test was given to each S. Analysis of variance and planned comparisons performed on the transfer data indicated that dimension preferences affected the learning rates on the transfer problem and that OL training altered the S's preferences among dimensions.

The results extend recent work in concept-shift behavior and provide support for the importance of dimensional preferences in children's concept learning.

TEACHING KINDERGARTEN CHILDREN TO APPLY CONCEPT-DEFINING RULES, *Evan R. Keislar* and *Samuel R. Schutz*, University of California, Los Angeles

In studies of concept learning, subjects are generally required to discover the concept-defining rule. In schools, however, children are usually given a rule and then expected to use it in identifying positive and negative instances. The purpose of the present study was to explore the value of instruction in applying conceptual rules in this "deductive" fashion.

Six children, 58 to 62 months of age, in a Children's Center, were individually given 8 daily programmed lessons, 8-12 minutes each. Six other children from a comparable population were given only the pre- and posttests.

For each deductive problem in the instruction and the criterion tests, 5 items involving the same rule were used. The child was told a story in which a missing object or person had to be found. The clue was given in the form of a description (i.e. defining rule) which the child then used as a basis for selecting 1 of 3 pictures which might be the missing object (the positive exemplar). The problems involved 4 of the most important types of conceptual rules (drawn from Bourne's list of 10): negation, conjunction, disjunction, and joint denial. An example of the latter would be: "Jimmy's sweater has no buttons and no sleeves." The pre and posttests of rule usage presented 8 problems, each with 5 items, for the 4 types of rules. In addition, an "inductive" test required the children to identify positive instances in a concept-identification problem without being told the rule.

The experimental group showed considerable improvement in applying each type of conceptual rule, gaining significantly more than the control ($p < .01$). They were also dramatically superior with the "inductive" problems, which required the use of new types of conceptual rules. Such evidence of transfer suggests the value of this type of training for young children.

THE EFFECTS OF MEMORY SUPPORT ON THE CONCEPT LEARNING OF TEST ANXIOUS CHILDREN, *F. L. Paulson,* Teaching Research Division, Oregon State System of Higher Education

Sieber and Kameya (1968) found that memory support improved the performance of test anxious children in a problem solving task. Their task required *S* to manipulate the materials physically, and to avoid certain moves that led to error. The present study was designed to determine whether this result was generalizable to a broader universe of behavior.

The performance of test anxious children in a concept learning task was investigated. The task 1) required no physical manipulation of materials, 2) elicited verbal guesses of the identity of the concept, and 3) provided no feedback regarding the correctness of *S*'s response. The concepts were represented as figural material printed on cards. 96 fifth and sixth graders were assigned to a 2x2 design combining Anxiety (High and Low) with Memory (Supported and Not Supported). Memory Support was provided by allowing *S* to view previously exposed exemplars.

ANOVA revealed Memory and Anxiety to be significant main effects. A planned comparison between the No Memory Support/High Anxious group and the remaining three groups confirmed the hypothesis that, without Memory Support, anxious *Ss* require more trials to criterion. The content of *S*'s verbal hypothesis was analyzed to determine if the groups differed in errors involving the compatibility of verbal hypothesis with information on the immediately preceding trial. Significantly more No Memory Support/High Anxious *Ss* made errors than *Ss* in any of the three remaining groups. The results indicate that the undesirable effects of anxiety may be partially overcome by reducing the requirements which the learning materials place upon the learner's memory.

TRAINING SEQUENCES WHICH FACILITATE CONCEPT LEARNING, *M. I. Charles E. Woodson,* University of California, Los Angeles

In a series of experiments investigating training sequences that facilitate concept learning, college students were trained on conjunctive, inclusive disjunctive, conditional, and biconditional conceptual rules in a reception paradigm. The problem contained two binary relevant attributes and two binary irrelevant attributes. A variety of training sequences were studied in several experiments.

Trial-to-trial redundancy of both the relevant and irrelevant attributes was found to facilitate concept learning.

Previous research has failed to distinguish between synchrony of relevant and irrelevant attribute change and trial-to-trial redundancy. Trial-to-trial redundancy was found to have the most significant effect.

Reasoning from the hypothesis that concept learning can be analyzed into two stages, an attribute identification stage followed by a rule learning stage, training sequences were constructed which facilitated attribute identification early in training and rule learning later in training. These were found to facilitate learning more than randomly generated control sequences or sequences which facilitated rule learning early in training and attribute identification late in learning.

These experiments indicate that the sequence of training instances is an important variable in concept learning.

Session 19.4 **Division C**

LEARNING FROM DISCOURSE: MATHEMAGENICS NO. 2

Joanna P. Williams, University of Pennsylvania, Chairman

THE INFLUENCE OF PARAGRAPH ORGANIZATION UPON THE LEVEL AND ORGANIZATION OF RECALL, *Lawrence T. Frase*, Bell Telephone Laboratories

The digits 313223121 are easier to learn if presented as 111222333 because they are organized or clustered together. Suppose that the digits represent sentences in a passage which refer to three different concepts. Coding sentences in this way allows one to express the conceptual organization of a passage or *S*'s recall protocols numerically, by determining the percentage of repetitions of conceptually related sentences. This study explored whether alternate organizations of a passage about chess would be equally efficient for learning and for constraining the organization of recall, and whether information about the conceptual structure of the passage would improve learning.

In order to evaluate chess play for monetary reward, 42 adults read a passage describing chessmen. *S*s read for 15 min., and were interrupted at 5 min. intervals for written recall. For Group N, each of 6 paragraphs described 8 attributes for one chessman. For Group A, each of 8 paragraphs described one attribute for 6 chessmen. For Group R (rote), the sentences were randomized. Half of the *S*s were told what attributes would be described and that there were six chessmen.

Group R recalled 37% fewer concept-attribute associations ($p < .005$) than Groups A and N (which did not differ), but Group R recall organization (79%) did not differ from Group A. Both Group A and R showed significantly lower recall organization than Group N (98%). Interrater reliability for organization scoring averaged .93. Information about the conceptual dimensions of the passage improved recall as more was learned (trials by pre-information interaction $p < .025$). Primacy effects were found only for organized passages.

The results indicate that readers may adapt to disorganized inputs and reproduce the information in an organized manner, but their learning will be retarded. Conceptual pre-information seems especially useful as the amount of internalized information increases.

EFFECT OF FREQUENCY AND POSITION OF QUESTIONS UPON LEARNING WRITTEN MATERIAL UNDER DIFFERENT INCENTIVE CONDITIONS, *Edward Patrick*, University of Massachusetts; *Lawrence T. Frase*, Bell Telephone Laboratories; *Harry Schumer*, University of Massachusetts

Previous studies have shown that learning from texts is improved if questions are inserted frequently after the material to which they relate. Prequestions, or infrequent postquestions, do not produce this effect. An important problem, aside from the general influence of incentive upon the learning of text, is whether high incentive motivation might override such stimulus manipulations.

270 college *S*s read a biographical passage in order to answer multiple-choice questions. Three groups expected to receive 0, 3, or 10 cents for each correct answer. Questions were placed in the text either before or

after the material to which they related, and were frequent (one question every 10 sentences) or infrequent (five questions every 50 sentences). Correct answers were not provided, and *S*s were not allowed to review the text. Incentive control groups read the text without questions. Dependent measures were correct responses on the 20 questions seen with the test (repeated), correct responses on 20 other questions not seen with the text (incidental), and time to complete the reading and test. Incentive influenced posttest scores ($p < .005$): 0¢<3¢<10¢. Question condition was also significant ($p < .01$): prequestion<control<postquestion. Time also increased with incentive ($p < .001$), and was influenced by questioning ($p < .01$): control<prequestion<postquestion. Analysis (controls excluded) revealed that scores on repeated questions were higher than incidental scores only if questions were frequent (interaction $p < .001$). Higher order interaction indicated that only when questions were infrequent did scores for the prequestion groups order themselves according to incentive conditions. Postquestion groups were resistant to these incentive effects on repeated questions but not on incidental.

These results indicate that as appropriate stimulus controls are removed, motivational factors begin to determine how much is learned from written materials.

EVIDENCE FOR A NEW INTERPRETATION OF THE ADDED PARTS EFFECT IN LEARNING FROM WRITTEN DISCOURSE, *Ernst Z. Rothkopf*, Bell Telephone Laboratories

In the added-parts method, new material is gradually added to previously studied portions of a written instructional document until it has been presented in its entirety. This technique has been found to be more efficient, under some conditions, than comparable whole or part techniques (Rothkopf, 1968). A new hypothesis about these findings is that (a) in the instructional sequences of the added-parts procedure, segments of new material tend to alternate with previously presented text segments; (b) such sequences favor the acquisition of new segments that just precede the familiar sections, probably because previously exposed text segments require relatively little information processing.

High school students ($N = 144$) familiarized themselves to the point of mastery with four attributive sentences about each of six trees in a fictional botany. Following this, *S*s were exposed to a string of nine sentences about each tree. Four of these were previously exposed sentences; five were new attributive sentences. The middle (i.e., 5th) sentence in each string was always new and was experimentally critical. Critical sentences (C) were always followed by two familiar (FF) and two unfamiliar sentences (UU). However, the sequence in some treatments was CUUFF while in others it was CFFUU. Training was experimenter-paced. Each sentence was exposed for ten seconds. Retention of the critical sentence was tested by the method of stimulated recall ten minutes after the completion of study.

As predicted, retention of *C* was approximately 15% better for CFFUU than for CUUFF presentation sequences. This result could not have been expected on the basis of ordinary interference theory, but requires a special sequential "microinterference" account. The present findings support the view that the observed advantage of the added-parts method was due to training sequences that alternated between familiar and unfamiliar text segments.

Session 19.6

INCIDENTAL LEARNING EFFECTS OF THE ADAPTIVE PROCESS INVOLVED IN FINDING INFORMATION IN A TEXT, *Francene Silbiger*, Fairleigh-Dickinson University, Madison; *Lawrence T. Frase*, Bell Telephone Laboratories

Only when sentences describing different attributes of a complex concept are separated in a text, does the concept name become critical for locating the related information. These criterion words allow the reader to decide whether to continue processing each sentence he encounters when looking for related information; thus they provide feedback which controls reading. Two studies investigated retention of these critical stimuli under differing methods of presentation and amounts of search.

In both studies *S*s were only told to find the name of a planet having certain attributes. In Study I ($N = 24$) there were 10 planets with 3 attributes each. Sentences were typed in list or paragraph form, and organized by name (no search involved) or by attributes (search involved). In Study II ($N = 61$) there were 15 planets with 5 attributes each. Groups had to search for 0, 3, 9, or 18 sentences. Presentation was only in prose format. The materials in both studies insured that a decision could not be made until all planets had been considered.

In Study I, groups which had to search recalled more names ($p < .005$). Other groups recalled more attributes ($p < .005$). List or paragraph format was not significant, and time taken to find the information was only suggestive. In Study II, retention of names increased with search ($p < .001$). Reading time also increased with search ($p < .001$), but amount learned per unit time was constant across search groups. Even names for which information did not have to be located were remembered when search was involved. Differences among groups and absolute level of retention showed little change after one month.

These results indicate that searching for information in a text has stable memory consequences for items which play a critical role in the maintenance of this feedback process, even though the reader may not intend to remember those items.

Session 19.6 **Division D**

MULTIVARIATE ANALYSIS

Marvin C. Alkin, University of California at Los Angeles, Chairman

MULTIVARIATE ANALYSIS OF VARIANCE FOR 2^k and 2^{k-m} DESIGNS, *Robert M. Pruzek*, State University of New York at Albany

In general, when a behavioral scientist states that he has conducted an analysis of variance, this means that he has conducted a *univariate* ANOVA. Surely the investigator may have studied both qualitative and quantitative factors, and he may even have included a number of covariates, but it remains likely that he has carried out his analysis on just one dependent variable at a time.

The above state of affairs is particularly unfortunate in the behavioral sciences, where each investigator is properly concerned with more than one dependent variable, or response measure, in virtually every study he conducts. In order to allow the simultaneous analysis of several response measures in a particular study, multivariate analysis of variance MANOVA has been developed (see esp., Roy and Gnanadesikan, *Ann. of Math. Stat.*, *30*, 1959). The chief virtue of MANOVA is that, under appropriate sampling assumptions, it allows for statistical inferences about the set of dependent variables, regardless of whatever non-zero correlations exist among them. While not widely applied as yet, and almost never examined in design textbooks, MANOVA can be used in nearly any situation where univariate designs are appropriate.

The major purposes of this paper are (1) to discuss a particular class of MANOVA designs, those called 2^k factorials, or 2^{k-m} fractional factorials, and (2) to show how computations for these multivariate designs can be conducted simply by generating various product-moment correlation coefficients. By thus restricting the range of designs to be considered, we can readily describe the basic distinguishing features of MANOVA, while at the same time facilitating simple computations work. We will show how conventional product-moment computer programs can be exploited in such applications, and examples will be given.

In the discussion, we indicate the relevance of certain classical methods, such as multiple and canonical correlation, discriminant function analysis, partial correlations, and factor analysis. We also indicate how MANOVA can be readily generalized to multivariate analysis of covariance (MANCOVA). Orthogonal and non-orthogonal analyses are contrasted. Finally, we consider possible implications of these designs for (1) analyzing trends using orthogonal polynomials, (2) making transformations of the dependent variables to simplify results, (3) using incomplete designs and controlling confounding for multivariate applications, (4) studying relationships between correlational and experimental methods generally, and (5) taking the reliabilities of dependent variables into account in analysis of variance studies.

INTERPRETING MULTIVARIATE ANALYSIS OF VARIANCE: A STUDY OF CURRICULUM INFLUENCES ON STUDENT SCIENCE INTERESTS, *S. Lloyd Newberry*, Armstrong State College; *Robert R. Rentz*, University of Georgia

Multivariate analysis of variance (MANOVA) provides a powerful technique for the analysis of educational experimental data when the data are multivariate. Unfortunately, a majority of producers of educational research lack the necessary training to make use of these relatively new analysis procedures. Scarcity of published research in education dealing with applications of MANOVA to experimental or quasi-experimental research designs creates difficulty for the typical producer of educational research when he is faced with an apparent multivariate problem. The difficulty lies not only in the choice of appropriate techniques, but especially in the subsequent interpretation of the results of these multivariate analyses.

Typical researchers in the substantive areas such as reading, curriculum, special education, learning, etc., have not the time nor inclination to wade through the necessary statistical sources in order to acquire a working knowledge of some multivariate techniques. It is relatively easy for the researcher to take his data to a computer center and request a particular analysis; however, interpretation of the results requires skill and understanding of those techniques. It seems that useful aids in the interpretation of MANOVA techniques would include examples with meaningful educational data. Cooley and Lohnes (1962), Jones (1966), and especially

Bock (1966) have provided excellent illustrations of MANOVA applications. It behooves the research methodologist to provide a larger literature on applications of newly developed methodologies if these advances are to benefit the research community.

This paper is concerned with the application of MANOVA to data involving a study of the effects of different kinds of science curricula on the science interests of elementary children. The focus is on interpretation of MANOVA results. The study involved comparing a conventional curriculum with a new curriculum (*Science—A Process Approach,* developed by AAAS) at grades four and six. The dependent variables were the nine subtests of a Science Interest Inventory developed by the senior author to measure the interests of elementary students in nine science areas (i.e. heat, light, sound, living matter, etc.).

MULTIVARIATE SEQUENTIAL ANALYSIS FOR EDUCATION, *Gerald J. Schluck,* The Florida State University

A multivariate sequential test for a hypothesis on the mean of a multivariate normal population is given, which generalizes the original univariate test given by Wald. Expressions for the average sampling number (A.S.N.) and power are presented. An alternative approach to the same hypothesis testing situation is considered—a group of vectors, rather than a single vector, is sampled. This alternative method has one advantage in that the procedure generalizes to the testing situation in which an estimate of the population dispersion matrix is used. If the population dispersion matrix is known, the grouped sequential test is a simple function of a chi-square statistic and its natural logarithm. If an estimate of the population dispersion matrix is used, the grouped sequential test is a simple function of an F (or Beta) statistic and its natural logarithm. Expectations do exist in both situations, and the expression for A.S.N. is given for the special cases in which either the null hypothesis or the alternative hypothesis is assumed to be true. An expression for the power of the test is not immediately available.

Monte-Carlo methods for a limited range of non-centrality parameters and for dimensionalities of two and three imply the economy in the multivariate testing situation is similar to that found in the univariate testing situation, namely, the expected sample size for the sequential test is approximately 40-50% of the fixed sample size test with the same error probabilities.

MAXIMUM LIKELIHOOD ESTIMATION OF COVARIANCE MATRICES FOR THE MULTIVARIATE RANDOM EFFECTS MODEL, *William H. Schmidt, William J. Bramble,* and *David E. Wiley,* University of Chicago

A general method is presented for estimating the variance-covariance matrices for the multivariate random effects model. An experimental design for which this model may be appropriate can be formulated as follows: 1) m factor levels are chosen at random from an infinite population of levels; 2) n experimental units are nested within each level; 3) p measures are obtained on each experimental unit.

The statistical model includes two parameter matrices—the within covariance matrix (Σ) and the between covariance matrix (Σ_b). Traditionally,

the estimation of these matrices has been accomplished by means of the multivariate analysis of variance. A problem associated with this method is that there is a non-zero probability that non-positive definite estimates of the two covariance matrices may result. The method proposed here eliminates this problem by employing the techniques of restricted maximum likelihood estimation. The results include maximum-likelihood estimates of the within and the between covariance matrices, and large-sample variances and covariances of the estimates.

The procedure involves defining the likelihood function and maximizing it with respect to Σ_b and Σ. This is accomplished by the method of Fletcher and Powell (1963) and Joreskog (1967). The restriction is made by defining Σ_b and Σ by their Cholesky factorizations, T_b and T, and maximizing with respect to them.

The implications of this method for educational research are discussed. One interesting educational application of this technique is to educational test data collected on pupils in a number of classrooms. The maximum likelihood estimates of covariance matrices may be used to isolate the between from the within classroom sources of variation and covariation influencing the test scores.

The procedure may also be extended to accomplish the simultaneous maximum likelihood factor analysis of the covariance matrices. A computer program for estimation of parameters is described. Data from the Louisville Twin Study (See Bock and Vandenberg, 1967) are used to illustrate the procedures.

Session 19.8 **Division D**

APPLICATION OF STUDENT RESPONSE MEASURES

Rodney W. Skager, University of California at Los Angeles, Chairman

A COMPARISON OF THE ABILITIES OF RECALL AND RECOGNITION ITEMS TO PREDICT HIGHER LEVEL COGNITIVE ACHIEVEMENTS, *Richard B. Smith*, Northern Illinois University

This paper represents an attempt to compare the abilities of recall and recognition items to predict performance on higher level *Taxonomy*-based science items. Item sets were constructed in which an attempt was made to hold content constant and systematically vary process according to the rationale of Bloom's *Taxonomy*. Each item set was constructed from a communication which consisted of the description of a science experiment, the resulting data, and the principle of the physical sciences which would result from the correct encoding of the data. The criterion set of higher level test items were items which required the subjects to "extrapolate" the principle, "apply" the principle, "analyze" either the experiment or a statement regarding the experiment, "synthesize" or determine the most logical next question to be asked or the next experiment to be performed, and finally evaluate the principle for use in a specific problem situation.

Procedure:

1. Administration of open-ended completion questions which required the student to recall the principles involved.
2. The presentation of the communication concerned with the same principles.
3. Administration of the multiple-choice item sets.

4. A comparison of the means and variances of two groups based on their performance on higher level *Taxonomy* items.
    a. Group I consisted of students who both recalled the principle and were able to answer multiple-choice items that required both a "knowledge" of the principle and an ability to "interpret" the principle.
    b. Group II consisted of students who were not able to recall the principle but after the communication were able to answer the multiple-choice items at the "knowledge of principle" and "interpretation" of principle levels.

Preliminary analysis of the data indicates that the students who have the ability to recall and recognize the principles do significantly better than students who can only recognize the principles on the higher level test items.

The finding, if further confirmed, would tend to support the notion that the delayed recall of relational principles is more indicative of a highly structured cognitive schema than is information at the recognition level. This is evidenced by a greater ability to "extrapolate," "apply," "analyze," formulate related hypotheses, and "evaluate" related information. The finding would also tend to suggest that we retain in a recallable form those materials that are related to more highly developed cognitive structures.

AN INVESTIGATION OF A TECHNIQUE FOR MEASURING TEACHER IMPACT THROUGH A PATTERN ANALYSIS OF STUDENT RESPONSES, *John W. Wick*, Northwestern University

The purpose of the research reported here was to investigate the usefulness of a new technique for measuring the impact of a teacher. With this technique, the impact of the teacher is gauged on the basis of output information, that is, all types of student responses. These responses were in the form of achievement or aptitude test items, as well as responses to interest or attitude inventories. The technique is unique with respect to measuring teacher effectiveness, in that it is a pattern analysis technique (Similar Response Analysis) which groups people together on the basis of the similarity of the pattern of their responses, and not on the basis of the number of "correct" or keyed responses.

The present study deals with responses given by upper elementary students after a full year with a given teacher. To the data accumulated for the students of a given class are added data from other students of the same age, randomly selected from other classes in the same school. The pattern analysis technique ordered the entire list so that students with similar responses appeared adjacent, or nearly adjacent, to one another. When the students of the single class were grouped at one end of the list, with the randomly selected students at the other, the most obvious cause appeared to be the impact of their one year's presence in the same classroom.

Some *post hoc* investigations of these classrooms so defined were included as a part of the overall study.

THE DEVELOPMENT AND INTERPRETATION OF CRITERION-REFERENCED TESTS, *Thomas Kriewall*, University of Wisconsin; *Edward Hirsch*, University of California at Los Angeles

Increased attention to the efficient facilitation of mastery learning through educational product research and improved management of learning situations has generated a renewed interest in the use of criterion-referenced tests. However, classical test theory has been inadequate to the task of guiding efforts in the design of CRT test items and the evaluation of results. This design and evaluation problem is viewed from a new perspective in this paper, providing an alternative to the classical test theory view.

A novel interpretation of item characteristic curves follows from the application of Wald's theory of sequential analysis. Suggestions are offered for reducing the testing time needed to detect mastery attainment levels consistent with the Neyman-Pearson theory of hypothesis testing. The relationship between mastery criteria and such sampling plans as single sampling, simple curtailed testing, and the use of the sequential probability ratio test is discussed. Applications are demonstrated in the area of computer generation and administration of criterion-referenced tests of mastery in selected arithmetic skills. The use of criterion-referenced test results, evaluated by sequential analytic techniques, indicates promise for reducing testing time and costs for specified behavioral objectives. This, in turn, promises the possibility of designing improved and extended capabilities for computer-assisted instructional management systems.

CONVERGENT AND DISCRIMINANT VALIDATION OF TEACHER, PEER, AND SELF-RATING MEASURES OF INTERPERSONAL RESPONSE TYPES, *Paul L. Wood*, University of Georgia; *Robert L. Akridge*, Louisiana Polytechnic Institute

This paper reports a study of the convergent and discriminant validity of teacher, peer, and self-report measures of three types of interpersonal behavior (aggressive, detached, and compliant) by the multitrait-multimethod procedure proposed by Campbell and Fiske.

The instruments used were (1) an Interpersonal Behavior Rating Scale, with which teachers rate their students according to aggressive, detached, and compliant responses to 12 hypothetical critical incidents, (2) a Sociometric Rating Scale, by which each member of a group or class rates every other member as being most like one of three behavior descriptions which are designed to represent aggressive, detached, and compliant behavior, and (3) the *Children's Personality Questionnaire*, a self-report measure from which an aggressive, detached, and compliant score was derived by combining scores of various CPQ factors. A step-wise multiple regression analysis was utilized to derive these composite scores.

Eighth sixth grade classes were included in the study: four from metropolitan Atlanta and four from non-metropolitan areas, including two from a small town (under 10,000 population) and two from a rural mountainous area. The subjects totaled 106 boys, 97 girls and 8 teachers.

The intercorrelations resulting when three different interpersonal response types were measured with three independent methods of measurement were examined. The reliabilities of the measures used were satisfactory, with the possible exception of the IBRS-Compliant scale and the CPQ-detached composite (.53 and .58 respectively, before correction for attenuation).

Considering the high degree of independence among the methods of measurement, convergent validity was demonstrated for the constructs of aggressive (.73, .35, .34) and detached types (.60, .28, .41) and moderate convergent validity (.35, .34, .36) was found for the construct of compliant types. Discriminant validity was demonstrated for all three variables. The three criteria suggested by Campbell and Fiske for discriminant validity were met to a notable degree.

PROGRAM PARTICIPANTS

Session No.

Abramson, Theodore ... 17.8
Ackerman, Margaret D. ... 1.9
Adams, Janice Freeman ... 6.3
Adcock, Carolyn ... 14.5
Adkins, Dorothy C. ... 17.2
Adler, Norman ... 7.3
Airasian, Peter W. ... 17.1
Akridge, Robert L. ... 19.8
Alexander, William M. ... 8.8
Aliotti, Nicholas C. ... 7.9
Alkin, Marvin C. ... 19.6
Allen, Anthony J. ... 9.3
Almy, Millie ... 8.8
Ames, Russell E., Jr. ... 15.3
Amidon, Edmund J. ... 18.5
Amidon, Peggy J. ... 18.5
Ammon, Mary Sue ... 8.9
Ammon, Paul R. ... 18.7
Anderson, Barry D. ... 7.3
Anderson, Bruce D. ... 14.8
Anderson, Charles C. ... 14.5
Anderson, Chris ... 11.5
Anderson, Dennis ... 5.7
Antonioni, David T. ... 5.11
Aquizap, Roman B. ... 10.3
Armstrong, Jenny R. ... 11.7
Arnoff, M. ... 5.10
Askov, Eunice N. ... 2.4
Austin, Stephen F. ... 19.1
Ausubel, David P. ... 15.5

Badel, Alden W. ... 14.14
Bagley, Clarence H. ... 12.13
Bahlke, Susan J. ... 7.4
Baird, J. Hugh ... 2.3
Baird, Leonard L. ... 12.13
Baker, Frank B. ... 10.3
Baker, Robert L. ... 7.5
Ballif, Bonnie L. ... 17.2
Barritt, Loren S. ... 7.6/17.5
Bartel, Nettie ... 17.5
Bauman, Daniel J. ... 5.7
Bavry, James L. ... 5.10
Bayer, Allen E. ... 12.3
Beaber, James D. ... 8.8/16.10
Beach, Mark ... 3.14
Beauchamp, George A. ... 7.1
Beggs, Donald L. ... 12.13
Bell, Norman T. ... 3.6
Bemis, Katherine A. ... 12.7/14.14
Bennett, Vernon ... 14.6
Bennik, Fred D. ... 3.6
Berger, Barbara ... 6.4
Bernstein, Allen L. ... 16.11
Berry, Carla F. ... 18.5

Session No.

Besag, Frank P. ... 17.4
Beyer, Barry K. ... 8.8
Biamonte, Aurora L. ... 13.5
Biesbrock, Edieann ... 1.5
Blaine, Daniel D. ... 18.10
Blanton, William E. ... 7.9
Bloland, Harland ... 17.3
Bloom, Richard ... 8.10
Blum, Abraham H. ... 14.5
Boersma, Frederic J. ... 14.14
Boger, Robert ... 6.4
Bolton, Dale L. ... 10.7
Bondi, Joseph ... 18.5
Bormuth, John R. ... 1.5
Borrowman, Merle L. ... 3.14
Bosco, James ... 14.7
Bouchard, Ellen ... 11.5
Bowers, Norman D. ... 8.8/12.12
Bracht, Glenn ... 18.11
Brady, Elizabeth H. ... 14.2
Bramble, William J. ... 11.8/19.6
Brandou, Julian ... 14.9
Braun, John R. ... 7.12
Brickner, Charlotte Ann ... 6.4
Brigman, Leellen ... 14.7
Briner, Conrad ... 7.3
Brophy, Jere Edward ... 6.4
Brown, Bob Burton ... 17.8
Brown, Janet R. ... 12.11
Brown, Laurence D. ... 8.10/13.5/15.3
Brown, Marsha D. ... 12.3
Brown, Robert D. ... 12.13
Brumbaugh, Robert B. ... 7.2
Bruno, James E. ... 10.2
Bryan, Miriam ... 2.4
Bryson, Juanita ... 1.5
Burmeister, Lou E. ... 12.5
Burach, Jonathan E. ... 6.7
Burhans, Clinton S., Jr. ... 8.8
Bussis, Anne M. ... 9.1
Button, Warren ... 6.7

Cahen, Leonard ... 4.6
Calfee, Robert C. ... 12.5
Carlow, Chester D. ... 11.7
Carey, Russell L. ... 19.1
Carlson, James E. ... 4.6
Carr, Julian ... 1.5
Carter, Alvin R. ... 14.6
Carter, Heather L. ... 17.1
Cartwright, Carol A. ... 12.7
Cartwright, G. Phillip ... 12.7/15.2
Cash, William L. ... 14.6
Centra, John ... 12.13
Chadbourn, Russell A. ... 12.12

Session No.

Chapman, Robin S. 8.9/12.5
Chittenden, Edward A. 14.14
Christie, Samuel G. 7.2
Cieri, Vincent P. 4.3
Clark, Michael C. 5.5
Coffin, Edwin C. 2.4
Cohen, Wallace M. 15.2
Cole, Henry P. 14.5
Colton, David L. 17.3
Connelly, F. Michael 7.1
Connett, William E. 9.3
Cook, Desmond 9.4/11.6/17.4
Cook, Harold 14.9
Cooper, James G. 3.6
Cope, Robert 12.13
Cote, Andre D. J. 14.5
Cottingham, Harold 14.6
Coughlan, Robert J. 8.1
Cowles, James DeWitt 1.9
Cox, Charlotte D. 3.7
Cox, David L. 18.7
Cox, Richard C. 18.10
Craig, Robert C. 15.5
Crawford, D. G. 1.10
Crawford, Jack 15.3/17.5
Crispin, David 19.2
Croft, John C. 10.7
Cross, Ray 14.6
Crossan, Donna 10.8
Curran, Thomas 18.7

Dacey, John S. 9.3
Dale, Irina 18.10
Damico, Sandra 11.6
Davidson, Robert E. 6.3
Davis, Gary A. 17.9
Davis, J. Kent 15.3
Davis, O. L., Jr. 8.15/9.2/16.5
Dembo, Myron H. 8.10
Demchik, Michael J. 16.5
Dick, Walter 3.6
Dillman, Duane H. 17.4
DiLorenzo, Louis T. 9.1
Dimitrosky, Lilly 8.8
Di Vesta, Francis J. 6.3
Doherty, Victor W. 17.8
Doke, Larry A. 1.9
Donaldson, William S. 1.10
Doppelt, Jerry 12.12
Doyle, Walter 18.5
Duff, William, Jr. 10.7
Dunham, J. L. 8.10/11.5
Dunn-Rankin, Peter 10.3
Durkin, Dolores 10.3

Egeland, Byron R. 12.13
Ehrenpreis, Walter 15.3
Eiszler, Charles 8.7

Session No.

Eledmann, Anne M. 16.10
Eller, William 8.7
Elias, Robert 9.3
Eliot, John 2.2
Ellis, Mancel R. 1.10
Ellison, Mary Lou 8.7
Engle, Patricia Lee 14.14
Elmer, Carolyn Lipton 7.3
Etzel, Barbara C. 14.1
Even, Alexander 7.9

Falusi, Arnold J. 10.7
Farley, Frank H. 1.10/4.5
Feldhausen, John F. 7.4/9.2/9.3
Feldman, David H. 3.6/8.8
Feldt, Leonard S. 11.8
Ferguson, Donald G. 8.8
Ferris, Donald R. 9.3
Findley, Warren G. 3.13
Finn, Jeremy D. 17.1
Fleming, Elyse S. 12.11
Flynn, John T. 2.2
Follman, John 7.9
Formanek, Ruth 1.16/9.1
Forbes, Dean W. 15.3
Foster, Julian F. S. 11.6
Fraenkel, Jack R. 7.1
Frase, Lawrence T. 19.4
Frederick, Wayne C. 7.5/8.10
Frederiksen, Norman 18.10
Freitag, Nancy 8.8/11.4
Friedman, Miles 9.1
Frobisher, Mary 9.1
Fromberg, Doris 8.8
Frost, Joe L. 16.3
Furst, Norma F. 7.7/16.10

Gaa, John P. 4.5
Gaite, A. J. H. 15.5/18.7
Garber, Herbert 4.5
Gavin, Alfred D. 8.15
George, Julio 8.1
Gill, Mohindra P. 7.9
Glass, Gene V. 3.7/5.7/11.8/12.12
Golub, Lester S. 8.8
Good, Thomas L. 12.4
Goodwin, William L. 10.8
Goolsby, Thomas M., Jr. 11.8
Gordon, Katherine K. 13.6
Gordon, Richard E. 13.6
Gotts, Edward 19.1
Grahlfs, F. Lincoln 7.4
Graves, Jack A. 18.7
Gray, William M. 16.11
Green, Bert F., Jr. 2.4
Greenberg, Selma 9.1
Greenough, Diane E. 18.6
Grennell, Robert L. 4.6

Session No.

Griffith, William S. 19.2
Grobman, Hulda 17.1
Guertin, Wilson H. 3.13
Gupta, Willa 9.1
Gustafson, H. W. 5.7
Guthrie, John T. 2.2

Hakstian, Ralph A. 7.12/12.12
Halasa, Ofelia 14.5
Halbrook, Mary Carol 13.5
Haller, Emil J. 7.3
Halpin, Andrew W. 8.1
Hampton, Sandra 4.6
Hansen, Duncan 3.6/4.3
Hapkiewicz, Walter G. 15.5
Harootunian, Berj 12.11/17.2
Harris, Margaret 3.13
Harrison, Barbara 8.8
Harrison, Forest I. 8.8
Hartnett, Rodney T. 11.6
Hastings, J. Thomas 5.10/17.1
Hatch, Evelyn 1.5
Haven, Elizabeth W. 8.8
Hawkes, Thomas H. 12.11
Hayes, Robert B. 8.7
Hedgbeth, James E. 15.2
Heinrich, Darlene 2.3
Helwig, Carl 8.1
Henderson, Ronald 9.1
Hendrix, Vernon L. 7.3
Hering, William M., Jr. 7.4
Hernandez, David 7.9
Hervey, Sarah D. 6.4/8.8
Hess, Robert 18.5
Hickrod, G. Alan 14.6
Hicks, E. Perry 8.8
Hill, Claude E. 19.3
Hill, Russell 7.7/8.15
Hill, Walter R. 12.5
Hiller, Jack H. 7.7
Hirsch, Edward 19.8
Ho, Wai-Ching 7.1/8.7
Hodgins, Audrey S. 10.3
Holder, David Lyal 16.6
Hofman, Richard J. 16.11
Holden, Marjorie H. 7.6
Hooker, Clifford P. 10.7
Hopkins, Kenneth D. 18.11
Houston, Samuel 10.7
Huebner, Robert 13.6
Hull, Ronald Edwin 10.8
Hunkins, Francis P. 11.4
Hurwitz, Emanuel 17.3
Hustuft, Dean H. 8.10
Hutchings, Lloyd 14.9

Inglis, Lloyd H. 14.7
Illenberg, Gregory J. 15.2

Session No.

Ilika, Joseph 16.3
Impellitteri, Joseph T. 5.11

Jacobs, Paul I. 13.5
Jacobson, Milton D. 8.8/18.11
Jason, Hilliard 9.3
Jensen, Larry 11.5
Jerman, Max 4.3
Jester, Robert E. 7.7
Johnson, David A. 18.10
Johnson, David W. 12.13
Johnson, Homer M. 8.1
Johnson, Irwin T. 19.2
Johnson, Joseph C., II 8.8/18.11
Johnson, B. Lamar 8.8
Johnson, Paul E. 18.7
Johnson, Ronald E. 11.5
Johnston, Donald P. 12.4
Jones, Reginald L. 14.7
Juola, Arvo E. 7.12
Justiz, Thomas B. 17.8

Kane, Robert B. 5.10
Kapel, David E. 12.11
Karnes, Merle B. 10.3
Karraker, R. J. 1.9
Kasier, Clarence J. 3.14
Katzenmeyer, Conrad C. 18.10
Kaya, Esin 2.3
Keesling, Ward J. 16.11
Keislar, Evan R. 3.8/19.3
Kenlee, Gilbert 2.2
Kennedy, George 2.2
Kennedy, John J. 9.4
Kennedy, James B. 8.1
Kershner, John R. 3.7
Kielsmeier, Cathy 14.7/17.5
Kight, Howard 17.2
Kirk, Barbara A. 12.13
Kirsner, Donald 14.2
Klausmeier, Herbert J. 5.10/15.3
Klein, Stephen P. 18.10
Kligensmith, John E. 9.4/14.7
Knemeyer, Manfred 5.5
Knutson, J. M. 11.5
Koehler, Roger 4.6
Koenke, Karl 8.7
Koff, Robert H. 12.11
Koopman, Peggy R. 18.6
Koran, Mary Lou 1.10
Kraft, Richard H. P. 7.2
Kreitlow, Burton W. 19.2
Kresh, Esther 2.4
Kriewall, Thomas 19.8
Kroeker, Leonard P. 11.8
Kunkel, Richard C. 8.8

Lamon, William Emile 14.9

	Session No.
LaNoue, George R.	7.3
Lazzaro, Thomas A.	12.13
Lehmann, Irvin J.	12.13
Leventhal, Eleanor	14.8
Liberty, Paul G., Jr.	7.2
Lindeman, Richard H.	13.6
Lindvall, C. Mauritz	5.5
Linn, Robert L.	12.13
Lipham, James M.	7.2
Lippert, Henry T.	3.6
Loadman, William E.	6.3/7.5
Locklin, Ralph H.	8.8
Loftus, Elizabeth F.	4.3
Longo, Alexander A.	4.3
Lundsteen, Sarah W.	2.2/16.5
Lutz, Frank	17.4
Lynch, William W.	10.8
Lyon, Judy	9.3
McCain, Floyd, Jr.	8.10
McClelland, Paul	8.8
McClelland, Samuel D.	7.9
McDonald, Frederick J.	1.10
McElhinney, James H.	8.8
McIntosh, Dean K.	15.2
McKeegan, Hugh F.	2.3
McKillip, William D.	11.7
McLaughlin, John A.	3.7
McLennan, Kenneth	14.6
McManus, John F.	4.6
McNeany, Robert H.	3.8
MacDonald, James B.	8.8
MacGinitie, Walter H.	7.6
Madaus, George F.	9.3
Madsen, David	3.14
Mahan, J. Maurice	3.8
Majer, Kenneth	4.3
Mallan, John	11.6
Mann, Lester	14.9
Manning, John	1.5
Manske, Mary E.	1.10
Marcus, R. Laverne	8.1
Marlowe, Byron H.	14.6
Massey, Randy	11.5
Massialas, Byron G.	8.8/11.4
Mauger, Sam	8.8
Mayeske, George W.	15.2
Maynard, Freddy J.	11.7
Medley, Donald	8.15
Meinke, Dean L.	15.3
Mendro, Robert L.	10.3
Menne, John W.	9.4
Merrifield, P. R.	5.10
Merrill, David	5.5
Mertie, Donald E.	16.6
Messuli, Jonathan C.	3.14
Metfessel, Newton S.	14.2
Meyer, Donald L.	16.11
Michael, William B.	14.2
Mietus, Walter S.	17.9
Milholland, John	16.11
Miller, Charles K.	3.7
Miller, William	7.9
Millett, Gregg B.	2.3
Minar, David W.	17.3
Moeller, Vernon	14.6
Monohan, William	14.6
Mooney, Robert F.	1.16
Moore, Jane	8.10
More, Arthur J.	12.7
More, J. William	3.8
Morrell, James E.	14.9
Moskow, Michael H.	14.6
Moss, Jerome Jr.	7.1
Mulkey, Gwendel D.	14.7
Murray, Frank B.	19.1
Mustico, Thomas W.	11.5
Muth, Rod	16.10
Naiman, Yossel	16.6
Natkin, Gerald L.	5.7
Neale, Daniel C.	12.13/14.8
Nealey, Stanley M.	7.12
Nelson, Jack L.	11.4
Nesbitt, Mary C.	13.6
Newberry, Lloyd S.	19.6
Nigro, George A.	12.12
Nitko, Anthony J.	10.3
Nord, Dennis	14.7
Norwood, C.	19.3
Novotney, Jerrold	8.1
Null, Eldon J.	8.1
Nunmedal, Susan G.	19.1
Nutting, Sue Ellis	11.7
Oakland, Thomas	17.5
Ober, Richard L.	18.5
O'Bryan, Kenneth G.	14.14
Ohnmacht, Dorothy C.	12.5
Ohnmacht, Fred W.	18.10
Okada, Tetsuo	15.2
Olson, David R.	10.8
O'Neil, Harold F.	4.3
Oppenlander, Lester	12.4
O'Reilly, Robert P.	15.2
Orme, Michael	9.2
Orost, Jean	9.1
Otto, Wayne	1.9/8.7
Packer, Athol	16.3
Page, Ellis B.	4.6
Palmer, Edward L.	15.3
Papanek, Miriam	9.1
Parker, Ronald K.	13.5
Patrick, Edward	19.4
Paulson, F. L.	19.3

Session No.

Paulus, Dieter 4.6
Personke, Carl R. 14.7
Persons, Edgar A. 16.3
Peterson, Richard E. 12.13
Peterson, R. Duane 19.2
Piele, Philip 12.4
Pierceson, David 1.5
Pikaart, Leonard 11.7
Pike, Lewis W. 7.6
Pingel, Louis A. 12.12
Planisek, R. J. 1.16
Popham, W. James 8.15
Porter, Andrew C. 12.12
Porter, Betty C. 6.3
Poulos, Nicholis 7.4
Powell, Evan R. 14.7
Preising, Paul P. 19.2
Prentice, Joan 1.5
Proger, Barton B. 14.9
Prokop, Manfred 18.5
Provus, Malcolm 16.6
Pruzek, Robert M. 19.6
Pucel, David J. 7.1/16.3
Pyatte, Jeff A. 7.1
Pyper, John Romney 5.7

Ragosta, Marjorie 17.4
Randall, Robert S. 10.2
Rankin, Richard 9.1
Raphaelson, Arnold H. 14.6
Reid, J. Christopher 3.13
Reilly, Anthony J. 18.11
Reller, Theodore 10.7
Rentz, Robert R. 19.6
Richards, James M., Jr. 5.10
Ringer, Wayne B. 19.2
Ripple, Richard E. 7.5
Rock, Donald 12.13
Romberg, Thomas A. 11.7
Rossmann, Jack E. 12.13
Roweton, William E. 17.9
Rusk, Bruce 10.3
Rystrom, Richard 17.5

Sabulao, Cesar M. 14.6
Salomon, Gavriel 17.2
Salter, Ruth 9.1
Samph, Thomas 12.4/17.8
Samuels, Jay S. 1.9
Sanders, James R. 10.8
Sassenrath, J. M. 17.2
Savicki, Victor 12.13
Scandura, Joseph M. 14.5
Schein, Judith 8.7
Schluck, Gerald J. 19.6
Schmidt, William H. 11.8/19.6
Schoenfeldt, Lyle F. 12.13
Schroeder, Glenn B. 12.7/14.14

Session No.

Schumer, Harry 12.13/19.4
Schutz, Samuel R. 19.3
Schwab, Lynne 11.4
Schwarcz, Robert M. 3.6
Scott, Norval C. 2.4
Scribner, Jay D. 7.2/17.3
Seaberg, John J., Jr. 7.2
Sears, Pauline S. 3.6
Seligman, Richard 5.10
Semmel, Melvyn I. 16.6/18.6
Senter, Donald R. 12.5
Severson, Herbert 4.5
Shea, Joseph 16.3
Shrable, Kenneth 17.2
Shuell, Thomas J. 3.8/15.5
Shulman, Lee S. 2.2
Sieber, Joan E. 14.14/17.2
Silberman, Harry F. 3.6
Silbiger, Francene 19.4
Simison, Diane 7.6
Sinclair, Robert L. 16.10
Singhal, Sushila 11.6
Sirotnik, Ken 11.8
Sitko, Merrill 13.6
Skager, Rodney 17.1/19.8
Slakter, Malcolm 4.6
Sloan, Jack 7.7
Smeraglio, Alfred 16.5
Smith, Brandon S. 7.1
Smith, Richard B. 19.8
Smith, Robert A. 16.11
Smith, Thomas M. 9.2
Smock, Charles D. 3.7
Smoot, B. R. 9.2
Snow, Richard E. 1.10
Snyder, Lester M. 5.11
Sokol, John 10.2
Spaights, Ernest 5.11
Speiss, Madeline 14.8
Spencer, Richard E. 18.7
Spielberger, Charles D. 4.3
Spring, Joel 6.7
Spuck, Dennis W. 16.10
Sroufe, Gerald E. 17.3
Stager, Mary 15.5
Stahler, Elizabeth 5.7
Stallings, Wm. M. 11.6/18.7
Stanfield, Robert E. 12.13
Starks, David D. 4.3/18.5
Start, K. B. 8.15
Stebbins, Linda B. 17.4
Steffe, Leslie P. 19.1
Stephens, Beth 3.7
Stern, Carolyn 1.5/3.8/9.1/11.4
Stoker, Howard W. 10.13
Stone, Douglas E. 16.6
Strickland, James F., Jr. 11.7
Stroh, Vickie 8.7

Session No.

Suppes, Patrick 4.3
Swanson, Richard A. 2.4
Sweeney, Jo A. 8.8/11.4

Tanaka, Masako 9.1
Tarte, Robert D. 1.10/7.6
Taylor, Raymond G., Jr. 14.9
Teager, Jayce 14.8
Terranova, Carmelo 4.6
Teska, James A. 10.3
Thiel, Richard 13.6
Thomas, Edmund D. 18.10
Thompson, D. L. 1.16
Tinsley, Drew C. 8.15/16.5
Tobias, Sigmund 7.5
Toole, David L. 5.7
Toole, Patrick F. 16.11
Toothaker, Larry E. 16.11
Tracz, George S. 10.2
Treffinger, Donald J. 7.5/9.3
Tseng, M. S. 1.16
Turner, Robert V. 16.10
Turney, Billy L. 14.7
Usdan, Michael D. 17.3
Utsey, Jordan 16.6

Vacc, Nicholas A. 18.6
Vandeventer, Mary 13.5
Van Mondfrans, Adrian P. 9.2/9.3
Vargas, Ernest A. 14.6/16.5
Veal, L. Ramon 1.5
Venezky, Richard L. 12.5
Vick, Marian Lee 18.11
Vogel, Francis X. 8.8

Walbesser, Henry H. 17.1
Walker, Decker F. 14.2

Session No.

Wallace, James 6.7
Wallen, Carl 16.6
Walls, Richard T. 6.3
Walster, William G. 11.8
Wang, Margaret 5.5
Ward, Beatrice 2.4
Ware, William B. 12.12
Warren, Thomas F. 17.9
Washington, Ernest D. 16.3
Waterland, Jean C. 5.11
Webb, Jeannine Nelson 17.8
Weener, Paul 7.6
Welty, Gordon A. 9.4
Wick, John W. 19.8
Wiesner, Carol 2.2
Wiggins, Thomas W. 8.1
Wiley, David E. 16.11/19.6
Williams, Joanna P. 1.9/19.4
Wittrock, Marle C. 19.3
Wolf, Richard 12.12
Wolf, Willevene 5.5
Wolfe, Lee R. 19.6
Womer, Frank B. 7.12
Wong, Martin R. 19.3
Wood, Paul L. 19.8
Woodbury, Charles A. 18.11
Woodson, M. I. Charles E. 19.3
Wright, Clifford J. 7.7
Wright, David J. 9.4

Yeatts, Pearline P. 13.6
Yelon, Stephen L. 14.7
Young, Beverly S. 19.1
Young, David B. 9.2
Young, Dorothy A. 9.2

Zahorik, John A. 14.7